among the constantly expanding number of potentially relevant articles and books just to find that nothing there "on my topic," the scholar is often able to assess quickly whether the argument and evidence of a piece are relevant to the scholar's current claims and whether they are consistent with them or challenge them. Even when sustained, the scholar usually has acquired supporting footnotes—though if the literature seems to say very much what the scholar wants to say, the scholar must then consider whether he has any residual dissatisfaction with the earlier literature, a number might, to KKV's final adumbration, that all research projects eventually be structured to permit "valid descriptive or causal inference." I have no objection to this requirement, depending on how it is interpreted. Critics of KKV have argued with some force that their book tends to equate "valid" inference with the adoption or imitation of particular methods of statistical inference (e.g., Collier, Brady, and Scarrow 2004). I agree that statistical methods have to be recognized as particular imperfect devices to implement inferential logic in relation to certain sorts of always-limited and flawed data—not as the logic of inference itself. In this poststructural age, we can make valid inferences without subjecting any data to any formal statistical analysis: we may not be sure who was really responsible for JFK's assassination, but we know bullets caused his death. In many circumstances, we can have extensive data sets that can be subjected to elaborate statistical analyses, but the data remain too flawed or biased to support valid inferences. But even if the process of reaching valid inferences should not be "the manifestation for methods or the process of it," it is one of the central points that things are supported by valid inferences is too narrow a fashion, but they should neither shrink from it nor fear it. Yet we still insist, in opposition to KKV but with a different emphasis, that through statistical inference is vital for social science, it remains only part of what doing good social science involves. It is at least as important—indeed, I think even more important to pursue research topics and make arguments that are substantively significant. Doing so is not easy, and it is true that the task of coming up with good ideas is partly intuitive. But the methods of more systematic introspection that have been so helpful to us, and by now we are familiar with them, can help scholars to discover and develop the good ideas that they already have within them. Hence I regard it as a qualitative method that has special value in generating richer fruit from our scholarly labors.

Note

1 Other writers may well have addressed these topics in ways I have not discovered. My credentials in this regard come chiefly from experience. I have supervised 29 P.A. dissertations and been a dissertation committee member on 52 others, including dissertations in sociology and American studies. Six of the dissertations I have supervised have won "best dissertation" awards from the American Political Science Association, in the fields of public law, women's studies, and community and other ethnic studies. One of these also won the best dissertation prize of the Law and Society Association. Seven of the theses on which I have been a dissertation committee member have also won ASPA dissertation awards, in public law, ethnic and agrarian politics, political philosophy, comparative politics, and federalism and inter-governmental relations. These dissertation committees have written a total of 34 articles that have been published in university press books to date, with several currently in production. Most of these dissertations have primarily or exclusively used non-quantitative methods.

References


Symposium: Multi-Method Work, Dispatches from the Front Lines

Introduction

Andrew Bennett
Georgetown University
bennettal@georgetown.edu

Multi-method approaches to research have generated considerable excitement in the field of political science in recent years. This is particularly true among graduate students, who are inspired by examples of excellent multi-method work by leading scholars, exhibited by their thesis committees to consider alternative approaches, and above all spurred by the apparent success on the job market of candidates whose research portfolio included multi-method work. Yet there are many challenges to using more than one method well, and there is a danger that graduate students in particular may risk becoming enticed by seemingly superficial claims that researchers will apply several methods poorly rather than doing one well, and that multi-method techniques will be back on to research problems for which they are not necessary or even useful. Perhaps most worrisome, compared to the shelves full of books and articles on one method or another, there is only a very small (albeit growing) methodological literature on how to combine different methods in the same research design (Lieberman 2005; Gerring and Seawright 2007). The emerging revolution in multi-method approaches has been driven not by methodologists, but by the practitioners of multi-method research, who have pioneered a diverse and innovative set of approaches and techniques. Methodologists are struggling to catch up and synthesize general lessons from the practices of researchers doing empirical multi-method work. This is the state of play. With the exception of myself as its editor (it), this consists of essays from those who have put multi-method research into practice, and it reflects their experiences from the front lines. My not-so-hidden agenda was to generate more evidence and insights for those of us engaged in multi-method work, offering general principles for its practice. To these ends, I asked each of the contributors to reflect on best practices in multi-method research, on their favorite examples of such research, and to conclude with a brief statement of where the field of multi-method research is at the moment. Although I did not ask the contributors, ranging from current graduate students to senior faculty, to focus on issues specific to their fields, there are contributors from each of the empirical subfields of political science: the contributors' works also represent a wide mix of different combinations of formal, statistical, and qualitative methods, although there are of course many possible combinations of methods—field work, ethnography, content analysis, surveys, case studies, formal modeling, simulations, archival analysis, interviews, and others—and not all of them could be included here. So, with due measure of the diversity of this research is that while the contributors noted many of their favorite examples of multi-method research, few contributions mentioned the same examples. Thus, the articles represented important contributions that are not necessarily representative of the wide range of multi-method research taking place, but they do give a diverse snapshot of the state of this research.

What emerges from this is a fair degree of consensus among the contributors on the prominence and difficulties of multi-method research. The authors were drawn to multi-method work by the potential for each method to offset some of the limitations of the others, a process that Thad Dunning calls "incremental matching." Scarcely any of the contributors had started with the intention of doing multi-method work for one reason or another, they were for the most part drawn to multi-method work by its methodologically important role by whatever methods they could muster. The essays convey the sense that each author emerged
satisfied that in the end they did indeed understand their phenomenon better through a variety of methods to bear. Yet the essays also embody a consensus on the chal-
gen of doing multi-method work as well. Such work can
be criticized from all corners or either the job market or by journal reviewers. The formal model or the statistical work might not be the very latest or most sophisticated that their respective
methodological communities have devised, the case study work might not cover all the relevant sources materials in all of the elaborate detail that a careful survey or experiment may not have anticipated all possible threats to validity, and so on.
There is also a danger that represents the dark side of triangula-
tion, and that is duplicitous. All three methods can compensate for the limitations of another, mistakes in any
one method can also cumulate when methods are applied se-
quently and build upon one another. Accordingly, several
contributors share Jason Wittenberg's view that mastery of
one method is better than more facility in several.
The bottom line from these practitioners is that despite the
considerable challenges and costs involved, multi-method research will remain a major feature of the high-level work.
If your interest is in this work it is my query to newsletter
readers for an essay from a recent or current graduate student
in multi-method work led to over a dozen responses. Rather
than choosing just one, asked Scottiegel, as the most
senior of this group, to be the lead author of a piece
co-authored by all of the respondents, and each contributor also
provided a brief synopsis of their thesis for this newsletter,
including a list of references of readers interested in
following up on their particular mixed-methods. The resulting
essay concisely captures the shared experiences and con-
cerns of this key constituency, noting the considerable pro-
blematic issues of coordination, the burden of data collection,
and the benefits that many of them received from attending training programs at Inter-University Consortium for Political and Social
Research (ICPSR), Empirical Implications of Theoretical Models (EITM), and the Institute for Qualitative and Multi-
Method Research (IQRMR). The authors also stress the need
for greater openness to multi-method work, especially such
work that includes a qualitative component, in the field's
leading journals.
The other essays in this symposium reinforce and build
upon these themes. That Dunning notes that multi-method research often involves numerous iterations among methods
rather than any simple linear progression from one method
to another. Some of these iterations are quick and intuitive, while
others are more deliberative, methodical, and deductive. He stresses in particular that the study of individual cases can
rarely inform the building of formal models. Daniel Carpen-
ter emphasizes this point as well, challenging the use of "as if" assumptions in formal models (p. 21), as an increasing
number of them appear to be doing, to inform their modeling from and test their models in qualitative case stud-
ies, rather than just using selected cases to illustrate models.
Sue Levin, drawing on her experiences as the author of more than two dozen articles (many of them multi-
method) in journals such as the American Political Science Review, the American Economic Review, International Or-
ganizations, and the American Journal of Politic-
ical Science, and the Journal of Conflict Resolution, focuses
on the problems of getting journal editors to pick appropriate
reviewers for multi-method work and getting department chairs and tenured professors to judge student work in these
areas as problems, Lohmann cogently argues that they are
urgent matters for promoting methodological cross-
fertilization in the field.
Jason Wittenberg, like several of the other contributors, highlights the importance of first developing a good question and
then choosing the methods that give the most leverage on it, rather than starting with methods and asking which
questions they can test. Wittenberg reminds us that the
"single county" study can often be disaggregated into
multiple case studies by comparisons across sub-units, over
time, or across cases. He illustrates this with his own research, in which his "single-case" study became three
hundred observations. Wittenberg concludes with an
especially useful summary of seven issues for scholars to con-
sider when contemplating a multi-method project or fieldwork in a developing country.
Finally, Hea Goemaness notes the importance of modeling and
explaining historically important individual cases even if
their detailed characteristics do not conform to the general
models of the field. This model of the field. His model stresses the importance of being careful with "off-the-shelf"
dataset codings, and he urges scholars to try coding them
several times in the cases in any off the shelf dataset they want to use in order the codings do or do not fit these
codes and concepts and purposes. Goemaness enjoins the users as well as the creators of such datasets to exercise responsibility and
provide feedback on codings. In this regard, he offers an ex-
perience in an important area of effort and experience
for datasets so that users can provide feedback on case
codings. This would require working through several diffi-
cult challenges—for example, if codings frequently change, re-
searchers will need to keep track of the codings operative on
the date on which they were accessed and perhaps periodi-
cally re-check their results as codings change. Also, dataset creators would have to decide whether to update codings frequently or to merely provide Web space for input from users that helps other users adjust their own case codings. Nonetheless, this offers a very promising approach to getting statistical and qualitative researchers to work together on is-
sues of common concern. Hopefully, conferences and work-
shops directed to this goal can be organized soon. This
proposal is of sufficient importance and magnitude that dataset
creators and data organizations like the Social Science
Bank need to pool their resources and work together to de-
velop suitable protocols for continuously improving dataset
codings (for an excellent listing of dozens of datasets on in-
national relations and politics, see http://ipumst-www.ssc.
ms/phenex/phenex.html).

In short, the growing emphasis on multi-method research is one of the most exciting and promising developments in a field that has for far too long been defined by isolated meth-
odological communities dealing at separate tables, but foster-
ing this development requires significant institutional changes to make it easier to carry out and reliable for all scholars, especially for the graduate students who are the field's future. Multi-method approaches are not for everyone, nor are they suited to every research puzzle, so we should not set unrealis-
tic expectations, especially when we consider graduate students about how common multi-method research approaches can or should be. Yet we need to make the field more hospitable for those who do aspire to the ambitious goal of carrying out multi-
method research. Journals need to find suitable reviewers for
multi-method work, and they may at times need to accommo-
date the higher word counts and/or Web-based appendices that such work can require. Departments need to ensure method-
ological pluralism in hiring and promotion, and to accom-
modate the fact that some research agendas are better suited to articles and other works to books. Departments also need to ensure that their graduate students have access to and re-
sources for cross-method training, either in-house or through
dedicated methods training programs. In these programs, in
need to incorporate multi-method approaches not just into their curricula but also into their cultures. Finally, organizations devoted to the infrastructure of the field, such as NSF and dataset providers, need to focus on getting the most out of their resources by offering seminars and classes, and by building in data, and perfectly captures the data-generating process. A slightly
different formal model would require a statistical estimator
of its own and might therefore produce significantly different re-
results—even for the same data. Now we might load a model for student—most models would admit. Sec-
ond, it assumes the model represents a pattern that regularly appears approximately. However, the strength of a formal model doesn't derive from its ability to explain a great many cases, as long as it can explain some (however important) cases that other mod-
els cannot explain. Formal models of war, for example, do not necessarily work as well as do principles and other simple
mechanism, only that the mechanism occurred at least once
and could occur again. In comparison to both the older and more recent statisti-
cal models, this case is a more fruitful and more suitable method to empirically examine formal models. First, case studies can trace the strategic interactions that form the basis of formal models. As shown in Schulz (2001) and Goeman (2000), case studies are the richest and most
informative of a wide range of choices considered and actions were taken, they can also show that some other actions were deliberately avoided in anticipa-
tion of the choices made by others. Moreover, in contrast,
over, case studies are not yoked to the assumption that any
unavoidably simplified formal model represents the true
data-generating process. Case studies can both recognize the
innovations and the strategic and political mechanisms. Case studies can trace and establish causal
mechanisms in the midst of a potentially overwhelming num-
ber of otherwise confounding factors. Even if the empirical
research does not exactly match the formal model, case
studies can often still offer a judgment of the relative fit and relevance
of the proposed mechanism. Second, formal models do not propose "covering laws" and do not claim universal and general generality. To show the relevance and power of a model, it suffices to empirically trace the causal mechanism in a handful of cases; one—preferably "miracle"—case. A statistical search for a general, statis-
tically significant pattern of a model's causal mechanism might reject the mechanism, even if the mechanism holds in a few substantively important instances. Thus, case studies can suffer from the uselessness of a particular model or only with a few cases. Moreover, some empirical patterns are (extremely) rare. Some rare events, however, are extremely im-
potent and deserve both theoretical and empirical attention. If there are only a few instances of a kind of impor-
tant event—say the decision to drop the atomic bomb, or
the disintegration of the Soviet Union—it is obviously impossible to use statistical methods to test any theory about those deci-
sions. This does not mean, however, that such a theory is un-
important or undesirable. I find it entirely plausible that forms
of rare events can provide important insights. The only
way to test such models is through careful qualitative
research and case studies. If qualitative research and case studies are natural part-
ners for formal models, they are essential partners for quanti-
tative research. Graduate students and faculty make enormous efforts to acquire the most up-to-date, most advanced toolboxes of statistical methods. Prominent journals publish relatively narrow articles that deal with some (minor) perceived statisti-
cal problems. Without question, these articles can be useful and extremely helpful. However, this no-
phratisticated technological arms race has gone on with little or no attention to the inputs: the actual data to be analyzed. Schol-
ars take data off-the-shelf with little thought or consider-
ation of the purposes for which the data were originally col-
lected, the coding schemes and decisions, and the reliability and accuracy of the data.

For example, the literature of War data, as well as the International Crisis Behavior and Militarized Interstate Dis-
putes data on conflict, record the outcome of the conflict, and all these datasets in essence record the military outcomes of con-
flict. Hence, the richness of the data is not the same as the broadness of the data. The military outcome of a conflict is a
nearly universal event; instead, such scholars want to know whether at the end of the conflict the loser or state was better off than before. Thus, the information content of the data is a perfect fit for the Egyptians, but the Egyptian crossing of the Suez Canal and overall military performance came as a shock to the Israe-
lis. As a result of the war, Sadat was able to make a deal with the Israelis that he could not have gotten before the war.

To make matters worse, most datasets come with a set of
coding rules, a bibliography, and a large matrix of numbers.
The lack of descriptive coding schemes then makes it
difficult for other scholars to understand the data at any
significant level of detail. We are simply to trust that a group (often) undergraduate students correctly implemented
the coding rules, although they may have little or no under-
standing of the relevant concepts and purposes of the
data.
Incentive and Selection Effects

This is where we run into a disconnect between what's good for political science and what's good for political scientists. Political science thrives best when it can rely on a mix of deep specialization and multi-methodism: the ideal is neither-or, but both. And yet there are powerful incentive and selection effects in political science that make the deeply specialized at the expense of the multi-methodist.

To see what type of scientist will dominate the political science landscape, we need to take a look at the hiring and promotion practices of political departments. The promotion to tenure in particular is critical. It depends on publications (especially publications in leading disciplinary journals), grants, and external letters of recommendations.

The following argument relies heavily on the assumption that journal articles count for tenure rather than books published with university press. In practice, the relative importance of articles and books varies across subfields and sub-subfields. Political scientists rely heavily on articles, political theory, on books, and comparative politics and international relations lie in between these two extremes. Within these subfields there is further variation, for example, scholars who study the U.S. Congress favor articles, those who study the U.S. presidency, books. Compared to an article, a book comes with more of an expectation that it will influence a phenomenon in full. The effect I describe below is thus more powerful for books than for articles.

For this reason, the bias in favor of deep specialization and against multi-methodism is likely to exist in muted form in subfields and sub-subfields that rely on books.

My argument also relies on the assumption that grants count for tenure. In practice, grants are less important in political science than they are in the natural sciences and medical sciences. Even so, they count for a little bit. Once again, we find variation within political science: the subfields and sub-subfields that are funded by foundations. People who study international relations, for example, rely on grants significantly more than those who study domestic politics or public policy.

Full professors with tenure carry graduate teaching and service assignments, including fewer publications in the flagship journal of the discipline, compared to a specialized political scientist, and not only because multi-method articles are generally longer and take more time to produce, but also because they take more time to review and resubmit and are less likely to get accepted.

My argument about articles carries over straightforwardly to National Science Foundation grants. On this dimension, too, while not a multi-method political scientist will be failing, for the National Science Foundation peer review system relies on deeply specialized peers.

The promotion to tenure also relies heavily on external letters of recommendation. A deeply specialized political scientist will get letters from his friends and family, who will describe exactly and at great length where he stands and how he fits in — and who like him, or at any rate, who need to get along with him because, independent of whether he gets tenure at this particular university, he is a member of their shared cluster for good. Meanwhile, the letters up short. They might still mention all over the place what they are written by political scientists who are specialized in all the various methods the candidate for tenure has employed. The specialized letter writers compare the multi-methodist to an overgrown kid who doesn’t know all that well in the first place — “to their own,” with the effect that the multi-methodist once again comes up short.

First, the faculty in the candidate’s academic department who vote on the tenure case might right see through the bias against multi-methods faculty and in their minds offset it. More likely they will not — after all, most of them are deeply specialized faculty who are comfortable deferring to the deeply specialized faculty charged with evaluating the multi-methodist.

Even if the faculty offset the bias in their minds, there is the problem of university-wide review committees and academic deans. They will judge the candidate based on how she looks on paper. Instead of looking at the candidate’s scholar- ship, they like to count the number of publications weighted by the number of pages and the number of citations of favorably and unfavorable tenure letters. The multi-methods candidate looks worse on paper than does a deeply specialized candidate, and so the bias against multi-methodism is further amplified.

To the extent that young political scientists anticipate the bias of the tenure process, they have powerful incentives to become deeply specialized rather than “spreading them- selves thin.” These incentives are self-reinforcing; the more single’s and the more other single’s are deeply specialized, all the more incentive will there be to remain deeply specialized and the multi-methodist. Quantitative methods are, of course, positively correlated: the bias of the American Political Science Review in favor of quantitative methods might actually be the result of a bias in favor of deep specialization.

Right off the bat, a multi-methodist political scientist will come up for tenure with fewer titles, fewer citations, including fewer publications in the flagship journal of the discipline, compared to a specialized political scientist, and not only because multi-method articles are generally longer and take more time to produce, but also because they take more time to review and resubmit and are less likely to get accepted.

Now suppose your employer employs a multi-method approach: you mix and match — say, game theory, a historical case study, and a regression analysis. The editors will assign a crew of reviewers, consisting of one game theoretician, one historian, and one statistician. These specialized reviewers will compare your methodological sophistication to what they are used to among their equally specialized colleagues, and you will fail. In fact, they will fail right after they get their process paper (after all, it can be a brilliant paper). More likely, if they don’t reject it outright, they will make revise-and-resubmit recommendations that are accompanied by incompatible demands, and then you will get a group of reviewers for revise-and-resubmits, and anytime you succeed in satisfying one reviewer you make another reviewer unhappy.

You will face another problem. In principle, political scientists who are specialized in a certain method might right see through the bias against multi-methods faculty and in their minds offset it. More likely they will not — after all, most of them are deeply specialized faculty who are comfortable deferring to the deeply specialized faculty charged with evaluating the multi-methodist.

Because of the reverse fundamental attribution error, it is in principle hard for multi-methodist political scientists like myself to make correct inferences about what is going on. I believe what I say in this essay is true, for three reasons:

First, I have published both single-method articles (e.g., Lohmann 1992, 1993, 1994b, 1997a) and multi-method articles (e.g., Lohmann 1994a, 1994c, 1998a), and so I can reexactly compare the submission process for the two kinds of articles. Getting published in a leading journal is hard, no matter who you are or what kind of research you do; but there is no question that it is considerably easier to decide to publish a multi-methods paper than it is to publish a narrowly specialized paper.

Second, I have submitted grant proposals for both single-method scholarship and multi-method scholarship. Over the years, the National Science Foundation has supported my single-method scholarship and declined to support my multi-method scholarship, which in turn was supported by foundations. This is clear evidence of the bias against multi-methods scholarship in the world. Multi-method scholarship, because of its holistic nature, is actually a better vehicle for raising money; but the grants that are considered prestigious in the minds of the scientific community — National Science Foundation grants — are elite.

Third, all of my promotions have been difficult, not prima- rily because of my multi-method scholarship (I have enough publications to show), but because my single-method scholarship cuts across two disciplines (political science and economics), and it turns out that my argument about the bias against multi-method scholarship holds up with a vengeance for interdisciplinary social scientists who are affiliated with a discipline-based department in the social sciences. The difficulty is measured, for example, by the time it takes for me to publish my peer reviewed papers (several years) compared to the standard time (less than a year). A couple of years ago, I came up for promotion to “Professor Step VI”; this title is granted by the University of California on evidence of great distinction, recognized nationally or internationally, in scholarship or creative achievement. My department promptly suffered a nervous breakdown and ended up voting to recommend to the personnel review, as a result of which my colleagues and economists drawn from other campuses of the University of California reviewed my case (after all, we are one university). The review, which lasted three years, is best described with the words “The horror! The horror!”

The point here is not to whine; after all, I survived, how-
ever baffled and battered, and one of the wonderful things about taking ecology is that after you have it, you are amazingly free to do whatever research you consider valuable.

To the Facilitator with his ambition, mindset, the protective parent, I would ask: Patience. In the case of Steven Levitt, an econometrician at the University of Chicago, recently Levitt penned an impossibly glib comment titled "Let's Just Get Rid of Tenure (Including Mine)" (2007). He calls upon the University to fill any tenured positions with someone who would be willing to go to a $15,000 salary increase. His best-selling book Freakonomics is subtitled "a rogue economist explores the hidden side of everything" (Levitt and Doherty 2005), but Levitt would hardly have been hired as an econometrician at the University of Chicago if he were truly a rogue economist; Chicago's economics department is famous for collecting the high priests of the economics discipline. If Levitt had turned into a rogue economist post-tenure, he would need the protections of tenure for sure. In fact, Levitt's ideas are utterly and diametrically, conventional, which is presumably why he is so very well rewarded--and rewarded as an econometrician at the University of Chicago if he were truly a rogue economist: Chicago's economics department is famous for collecting the high priests of the economics discipline. If Levitt had turned into a rogue economist post-tenure, he would need the protections of tenure for sure. In fact, Levitt's ideas are utterly and diametrically, conventional, which is presumably why he is so very well rewarded--and rewarded as an economi

The first multi-methodists tend to use different eclectic mixes of methods and hang around in different combinations of specialized clusters, and they don't generally know each other as well as specialists know each other in other words, multi-methodists don't naturally flock together the way specialists do. Second, if multi-methodists were to cluster together, multi-methodism would lose much of its power in correcting the dualism of disciplinary paradigms. After all, the whole point of a multi-methodist is to exist at the fringes of deeply specialized clusters and connect them. Moreover, if multi-methodists were to form a group of their own, the more practical goal might be to redefine the methodological cutting edge. To some degree, this is what happens to political scientists who are embedded in professional schools (public policy, public health, business, and so on) rather than discipline-based departments of political science. To the extent that they hang out with scientists who are affiliated with professional schools, they lose their connections to the deeply specialized political scientists, and it shows. There is no obvious structural solution to the problem of the bias against multi-methodological political scientists. All we can do, it seems, is to rely on the education and goodwill of key decision-makers, such as journal editors and foundation ofﬁcers and department chairs and academic deans. This is a winn


Lohmann, Susanne and Sharyn O'Halloran. 1994c. "Divided Govem-


Lohmann, Susanne. 1984b. "Fearfulness and Central Bank Indepen-


Dealing with the stochasticity of predictions, especially in complex systems like the climate, requires robust methods to quantify uncertainties. Techniques such as ensemble forecasting and Bayesian inference are increasingly employed to assess the reliability of model outputs. These approaches not only provide estimates of the model's predictive uncertainty but also enable the incorporation of prior knowledge and expert judgment. By doing so, they enhance the decision-making process in fields ranging from renewable energy planning to disaster management.

For instance, in renewable energy planning, models are used to predict the impact of large-scale wind and solar farms on local ecosystems. However, these models have inherent uncertainties due to factors such as weather variability and species distribution. By employing ensemble forecasting, decision-makers can better understand the range of possible outcomes, which is crucial for planning resilient and sustainable infrastructure.

In disaster management, models are used to predict the path of tropical cyclones or wildfires. These models are inherently stochastic due to the chaotic nature of these events. By using Bayesian inference, decision-makers can update their predictions as new information becomes available, leading to more effective evacuation and response strategies.

Moreover, these methods are being integrated into decision support systems, allowing policymakers and stakeholders to engage in a structured dialogue about the uncertainties inherent in climate change projections. This not only helps in setting realistic targets but also in designing flexible and adaptive policies that can respond to different future scenarios.

In summary, robust methods to deal with the stochasticity of predictions are essential for making informed decisions in a world characterized by increasing uncertainties. By fostering a culture of embracing uncertainty, we can make better decisions and build more resilient societies.
experience, search committees prefer candidates with mastery of one method to multiple methods. Poorly executed research may end up pleasing no one.

This essay is an attempt to demystify the practice of multi-method research by illustrating how I executed the project that resulted in my master's thesis. My work, Crucibles of Political Loyalty: Church Institutions and Electoral Continuity in Hungary (Cambridge University Press, 2006), was the largest audience of graduate students and others who want to employ multi-method research. Others may be intimidated by the daunting and frustrating. As will become clear further below, there is no magic bullet. The practice of multi-method research involves options not taken, difficult tradeoffs, and a willingness to make midcourse improvisations in response to the new data arising. The bulk of the essay will elaborate the reasons for my choices and how my dissertation and book have been received. I conclude with some advice for graduate students on the peril and promise of multi-method research.

The Dissertation I Might Have Written But Didn't

The idea for my dissertation came from observing a peculiar feature of post-communist Hungarian politics: the emergence of political parties with the same names and slogans as pre-communist parties. Why should old symbols and labels reappear and gain electoral traction after four decades of communist rule? After more research, I realized that similar partisan continuities appeared elsewhere in Eastern Europe and, in a different form, in newly democratized countries of Southern Europe and Eastern Europe. I was very excited about the possibility of a large-N analysis of all countries where democracy was interrupted by some period of authoritarian rule. I might have collected data on the duration and nature of authoritarian rule, party systems, opposition behavior, and many other potential predictors of political continuity with the pre-authoritarian past. Part of my thesis would have consisted of cross-national statistical analyses. I would then have elaborated the detailed workings of the argument through carefully selected case studies.

I do not recall ever having seriously considered this possibility (though it still strikes me as an excellent topic: you would have a ton of time). Communist politics was politically divided, and my tacit assumption was that cross-regional comparisons, while technically possible, were of limited analytic utility. Analysis of authoritarian and Southern European authoritarian periods seemed too different from Eastern European communism to permit meaningful comparison. Moreover, I had been trained to believe that one could not understand a country's politics except through mastering its language and immersing oneself in the society, typically through at least a year of fieldwork. Consequently, although I was no stranger to statistical analysis, the idea of serious research in more than one country seemed daunting. As an uncertain value. I might have been swayed otherwise if someone had argued that my job prospects depended on it, but post-communist candidates seemed to be getting jobs, the market was distant, and one country seemed enough. Quite possibly, this shortcoming might pose problems of generalizability, but I came to realize that the only way to address the larger puzzle of long-term political change was by empirical examination. In short, it was a particularly difficult case. The bulk of my empirical research would be limited to Hungary.

The Dissertation I Did Write

I began the project without any explicit intention of employing multiple methods. If I had to characterize how my thesis (and later book) came to have its particular blend of qualitative and quantitative analysis, I would have to say that it did what seemed most useful for answering the question. This is not meant to be glib. I was no less interested than contemporary students in doing good work. The question was on the agenda rather than on the methodological eclecticism per se. That nonetheless ended up employing an array of tools is evidence that then-existing folks wisdom and common sense or how to do good social science often entailed the use of mixed methods. This is not to imply that for every question there is an obvious research design. It is all too easy to err. Rather, it is an acknowledgement that many researchers were employing multiple methods, at least in some cases, long before they achieved their current exalted status.

Why should old patterns of mass political loyalty reemerge after prolonged economic, social, and political disruption? My research strategy tried to gain as much leverage on this question as possible within the constraint of focusing on a single country. I pursued a three-pronged approach, each part of which was designed to address a different anticipated dimension of the thesis. The first two subunits of the book, for instance, was the charge that Hungary was not an interesting place to explore the question. The "why did you study country-case" question is among the most common ones, and we owe unto whom? ever such a question is a reasonable one. I spent a lot of time, notably in 1996/1997, when I was writing my book, to try to figure out how to make Hungary an "interesting" country. I did so by focusing on the process of liberalization and the role of the churches in the transition to democracy. I then used this to frame the question of political continuity. Prior theory tended to focus on the less disruptive authoritarian systems of Latin America or Southern Europe, where the churches were significant actors. I understood that opposition parties opposed to the dictatorship were invited to account for partisan persistence. Under communist civil society was far more comprehensively destroyed or co-opted, and could not perform the same function. Thus, whatever was producing continuity in Hungary had to be different from what was causing similar outcomes elsewhere. The advantage of studying Hungary, then, lay in the potential for exposing a new transmission mechanism.

The second and related problem to avoid was what King, Keohane, and Verba (1994: 208) refers to as the "p=1 problem." One national-level observation of continuity yields precious little inferential leverage. To counter this I disaggregated the dependent variable. There had been studies of regional electoral continuity in Hungary, but changes in internal borders rendered the second imperfect. A lot has been written about the country's history claimed that my findings merely reflected the fact that the region I had focused on had always been among the most conservative and Catholic parts of Hungary. Thus, we did not question the hypothesis of Hungary as a whole. Quantitatively people questioned some model specifications and my reliance on ecological data. I took these criticisms very seriously because it felt like if I didn't have robust results I would have wasted my time. The research had not yielded significant to increasing confidence in my findings. On the qualitative side, I replicated the archival research in a predominantly Protestant region. This showed that my initial results were not a fluke and provided leverage on cross- regional effects that I could not explore with materials that focused only on Catholic activity. On the quantitative side, I replicated a number of the findings in other settings that might illustrate the process. There was no guarantee of access to such materials even existed. The general dearth of information made such qualitative piece by piece for the most challenging part of the project. I spent months of my second extended period of field research exploring provincial archives, where I discovered that the survival of right-wing attachments was rooted in the successful efforts of Catholic parish priests to preserve local church institutional continuity. We are conditioned to think of archival materials as inherently qualitative. In this case, however, they yielded invaluable local-level data on mass loyalty to the churches. I was thus able to demonstrate clerical influence both quantitatively, for a smaller sample of settlements, and qualitatively through interpretive analysis of Communist Party and church reports.

From Dissertation to Book

I received many suggestions for improvement as I endured the job market and prepared the book. One of the more common was that I should add another post-authoritarian case. There were three main factors that led me to publish the book as it was. I could see no logically compelling reason for the considerable extra effort. Although there was certainly a payoff to knowing that the basic argument held up in a different political context, in the end the primary unit of analysis was locality, not country. Including settlements from a different political party would not add variance that did not already exist within Hungary. Another frequent suggestion was that I include an in-depth narrative of how the struggle between parish priests and local party cadres played out in a single village. I seriously considered this because it would have improved the argument's rhetorical force, but I decided I was better off leaving the book as it was. This was an easy call: I did not have enough space to work on the narrative. I decided I could do without the in-depth narrative with the strategy of the book. The book was thus not representative of my work as a whole. Quantitative people questioned some model specifications and my reliance on ecological data. I took these criticisms very seriously because it felt like if I didn't have robust results I would have wasted my time. The research had not yielded significant to increasing confidence in my findings. On the qualitative side, I replicated the archival research in a predominantly Protestant region. This showed that my initial results were not a fluke and provided leverage on cross-regional effects that I could not explore with materials that focused only on Catholic activity. On the quantitative side, I replicated a number of the findings in other settings that might illustrate the process. There was no guarantee of access to such materials even existed. The general dearth of information made such qualitative piece by piece for the most challenging part of the project. I spent months of my second extended period of field research exploring provincial archives, where I discovered that the survival of right-wing attachments was rooted in the successful efforts of Catholic parish priests to preserve local church institutional continuity. We are conditioned to think of archival materials as inherently qualitative. In this case, however, they yielded invaluable local-level data on mass loyalty to the churches. I was thus able to demonstrate clerical influence both quantitatively, for a smaller sample of settlements, and qualitatively through interpretive analysis of Communist Party and church reports.

Advice for Graduate Students

My experience may not be wholly representative, but I do think it offers a few lessons for those contemplating or already engaged in mixed-method research.

1. Choose a question, then a method. It sounds obvious, but the availability of automated tools allows us to generate output even in the absence of a research question. Resist the temptation to number methods before nailing down the purpose of the analysis.

2. If your research is primarily on one country, make sure you get that country right and are prepared to defend your choice. Cross-national research on a single country will have equal mastery over their cases, even those they investigate more thoroughly than a nested design.

3. If your research is primarily on one country, make sure you have sufficient within-country variation between subunits, over time periods, or across functional issue areas. Make sure others know that the unit of analysis is not simply the country. Correct those who dismiss your work as a "case study".

4. Be prepared to get hit from all methodological sides. Good departments will expect you to use all your methods equally well.

5. If you work on developing countries, do not assume that others appreciate the difficulties of data collection. People who google their data may require special enlightenment.

6. If you work on developing countries, do not expect forgiveness for lacking the kind of data that are available to those who address similar questions in developed countries. People who have done fieldwork will sympathize with your plight, but others may penalize you for taking a question that could not be fully answered.

7. If you work on developing countries, do not expect much extra credit for overcoming obstacles to data collection. Those who have done fieldwork will laud your ingenuity, but in the end good departments are more interested in what you have done with the data than in the data themselves.

8. I conclude by re-emphasizing the importance of starting with the right research question. In the final analysis research is truly competely, you may be forgiven some minor sins, but no amount of methodological razzle-dazzle can compensate for a poorly posed problem.
The Role of Iteration in Multi-Method Research

Thad Dunning
Yale University
thad.dunning@yale.edu

Self-consciously “multi-method” research seems on the rise in many corners of the discipline. Recent political science dissertations, in particular, seem to draw increasingly on some combination of fieldwork, game theory, statistical analysis, qualitative-historical-institutional comparisons, ethnography, and other approaches.

Why is multi-method work so attractive? One powerful reason is that such research allows the possibility of triangulating across a great variety of methods, allowing scholars to leverage the distinctive but complementary strengths of different research methods to make progress on substantively important topics. Thus analysts strive to move beyond evidence on aggregate correlations and evidence on mechanisms, to combine broad general theory with fine-grained detail from case studies, to motivate a large-N analysis by a few well-chosen cases, or to marry “data set observations” to “casual process observations” drawn from focused qualitative research (Collier, Brady, and Sears, 2004). The particular ways in which different methods should or could be combined, however, has remained the subject of debate (Laitin, 2002). For one, in multi-method work there always remains the possibility that we will get things wrong three ways (two or four) instead of just one. A statistical analysis might suggest that the likelihood of this occurring diminishes in the number of methods: if each method represents an independent approximation of the truth, the precision with which we estimate this “truth” should increase as the number of methods grows and sampling error diminishes. From this perspective, an N of three or four, where the N is the number of methods, should be at least a little better than an N of one.

This statistical analysis seems misleading, however, because applying different methods is not like drawing balls independently from an urn. In good multi-method work, various commentators methods are supposed to inform one another. Then “draws” from the methodological urn, rather than being independent, may instead exhibit strong dependence. At least in principle, adding a new method to a multi-method study could conceivably exacerbate rather than ameliorate the flaws of each of the others.

The dependence of each new methodological “draw” on prior methodological choices may be one reason that some methodologists encourage documenting the process by which scholars go about multi-method work—for instance, describing the order in which various methods were used or applied (Bennett and Brunnecoller, 2004). Yet if one event always affects where one ends up, the Pandora’s Box problem is not quite a Polya urn. In a typical illustration of a “Polya urn process,” a ball is drawn at random from an urn filled with two balls of different colors. If a red ball is drawn, then an additional ball of the same color is returned to the urn, and the procedure is then repeated a large number of times. As Pierson (2006: 253) and others have emphasized in analogies to path-dependent processes in politics, in such a process the initial sequence of at-random draws matters greatly for the ultimate distribution of balls in the urn. In addition, the ultimate outcome of any particular trial (i.e., any “large sequence of draws”) is ex ante unpredictable. They also indicate that even a trial with an urn filled with balls mostly of one color or the other.

This Polya urn analogy, as applied to multi-method research, seems too pessimistic. For one, in the iteration between various methods there are often ample opportunities for cross-method correction and revision. For another, even in the elaboration of any “simple” method, the characteristic features of research may still be relatively new and can have an important role. In this way, the idea that analysts “apply” one method and then exploit another may not characterize all multi-method research. The central issue therefore remains exactly whether our methods can be combined in a way that can generate a “multi-dimensional conspiracy” (with apologies to Albert O. Hirschman) in favor of scholarly progress. In this essay, I offer just a few thoughts in this vein, drawn from well-chosen examples from recent multi-method research. Several authors have recently discussed the ways in which case studies and large-N analysis can inform and complement one another (e.g., Lieberman 2005; Gerring and Sears, 2007), but there have been perhaps somewhat less sustained attention to the relationship between game-theoretic formal models and other methods.

I seek to make two simple points. First, I discuss the ways in which building an applied formal model—apparently an eminently “deductive” exercise—may in fact involve interventions and especially modes of concept formation usually more closely associated with other methodological approaches, including qualitative” methods. Second, in discussing the relationship between models and case-study evidence, I briefly reflect on the challenges associated with what Skocpol and Somers (1980) called “incongruence and the demonstration of theory.” In both cases, my emphasis is on how formal models and other methods may inform each other in ways that are more intuitive and even seamless than the image of sequential “draws” from a methodological urn would suggest.

Models, Concepts, and Cases

To pick an example not quite completely at random, and with my thanks to Peter Ordeshook (and imagination), I illustrate these points with a discussion of some of my own recent work on the impact of natural resource wealth on political regimes (Dunning 2007). It may be useful to briefly describe the overall trajectory of my research before exploring several issues and challenges that arose in the course of conducting it. At a near-consensus has emerged among scholars working in this area that oil and similar natural resources promote authoritarianism (Beijing, 2004) and that resource-rich (but not resource-dependent) countries in the world are liberal democracies, while a somewhat older case-study literature has suggested that oil historically promoted democracy in Venezuela—strong in Latin America’s most stable democracies for several decades in the second half of the twentieth century. My research was inspired both by the observation of an apparent contradiction between literacy and culture and my familiarity with this several anamalous cases.

At the time I began this research, my disquiet about the chain of oil and hindered democracy was also motivated by my study of recent game-theoretic work on the influence of redistributive conflict on the emergence and persistence of democracy (e.g., Acemoglu and Robinson 2006). If resources really shaped the fiscal basis of states in the way the literature on “rentier states” suggests—that is, if oil and other resources displaces non-resource taxation—then I think models one might expect more mixed effects of resource wealth, since resources might help the redistributive pressures democracy may sometimes be able to resist the idea that resource revenue could thus have mixed effects on the regime type matched intuitions I had drawn from cases where I had done brief initial field visits, such as Botswana, Chile, and Venezuela.

In conjunction with the literature on the politics of rentier states and with further fieldwork, I began to develop a game-theoretic model to help analyze these issues. There are always many analytic choices that go into the specification of competitive models. In the case of my research, for instance, should resource rents appear as a term in the government budget constraint, or in a function giving the wealth or income of different societal actors (or both)? In the model’s underestimation of economic structure, what should be the relationship between sources of rent to the non-rent-source sectors of the economy? These are just a few of the important questions that had to be addressed in this paper, so I avoid countervail or its equilibria analysis.

First, I think it is important that we know how knowledgeable case studies, the previous literature, and other sources of prior information can inform answers to such questions; a bevy of “multi-method” approaches may play a crucial role in helping to motivate and inform the structure of a given formal analysis. In my own case, the previous literature provided some helpful guidance on the analytic choices mentioned above. For example, the “Rostow states” suggested that resources should like oil tend to provide a ready source of government revenue and also to flow into the fiscal coffers of the state (i.e., the government budget constraint) like “manna from heaven,” without substantial intermediation of numerous societal actors. Following this logic, resources should appear only in the government budget constraint of the model and not in a function giving the (pre-transfer) income of private citizens, at least as an initial matter. For another, the work of Hirschman and others had long suggested that “enclave” natural resource sectors developed weak interdependencies, and “forward” and “backward” linkages to non-resource economic sectors. This idea suggested that resource and non-resource economic sectors might plausibly be modeled as independent, linked only through the channel of government spending. I found that a key to developing a useful applied model was to find ways of formalizing the contrast between the two sectors, but without giving further credence to the claims of the rentier state literature (Dunning 2007, Chapter 5).

These broad efforts may go to suggest that the process of developing a game-theoretic model can itself be a “mix-method” process. Because analysts may draw on well-developed concepts and previous results in the field to stipulate core assumptions, developing a model may be considered a process that is both backward and forward looking. It might also be argued that, as the examples, as mentioned above, that some of the distinctive strengths of “qualitative” methods, including especially tools for concept formation, can and should inform the development of applied models.

However, this discussion raises the important issue of how to evaluate model-derived hypotheses empirically and, more generally, the relationship of models to various forms of empirical testing. The classic “rival” models are the UV. Most analysts rely on “data sets” to evaluate these models, it might be argued that any empirical test is strictly an empirical test of a hypothesis. If we consider the statistical methods used in recent research as the standard, it would seem perhaps somewhat less sustained attention to the relationship between game-theoretic formal models and other methods.

Such “parallel” strategies should probably be an important part of evaluating a theoretical model, formal or not. Theories have observable implications, and at least a necessary condition for a valid theory should be that those implications tend, in fact, to be observed where the theory says they should be. Yet such parallel demonstrations can also be unsatisfying, for reasons of both methodology and theory. Skocpol and Somers suggest cases can end up seeming simply a way of underscoring the “plausibility” of a theory, its ability to “order the evidence” without, however, helping to refine or push the theory forward.

Analysts might strive for a more fruitful marriage of formal and empirical, particularly case-study, research in several ways.
Trends in Multi-Method Research: Sailing Ahead, Reckoning with Old Risks and New

Scott Siegel with Ariel Araham, Julia Azari, Ashwin Chhatre, Bridget Coggin, Jana Grittersova, Matthew Ingram, Matthew Lieber, Claire Metelis, Tom Pepinsky, Andrew Pierper, Karthika Saksikumar, and Prerna Singh

As Skocpol and Sonenshein (1980: 191-2) also emphasize, the parallel demonstration that avoids the "replication mistake" (in which the same theory is simply applied to multiple cases) when a theory predicts different outcomes across different cases—i.e., when the cases help to elucidate what a formal model would fail to tell—constitutes the "comparative statics" of a model. Evaluating these comparative statics through analysis of new cases that did not originally motivate the work, or new within-case evidence drawn from cases that did, can also provide an important view of the strengths and weaknesses of the theoretical models empirically. Another point is that for those oriented towards formal work, case studies cannot not only provide evidence on the observable implications of a theory but can also help to create new models, an advantage of case studies that I found especially useful in my own work (Dunning, 2007, Chapter 7).

An ongoing tension between methods that thus better characterize multi-method work does than the idea of one methodological “flow” followed by another. If one finding or methodological approach does condition the next, the multi-method research hardly reproduces the mythology of Polya’s prose process. Indeed, the strengths of different methods may inform each other at every stage of the research process, serving to balance and correct each other. It may therefore be worth reflecting on how apparently disconnected research efforts, such as concept analysis and genetic-geologic modeling, may in fact complement each other in useful and possibly unexpected ways.

References

Bennett, Andrew and Brian Brunowmoff. 2006. "Where the Model Frequentely Meets the Road: Combining Statistical, Formal, and Case Study Methods." Unpublished manuscript.
issue to which both the discipline and policymakers demand answers. Although the problem of avoiding policy-relevant issues is of concern for the study of multi-method research, solving this multi-method research to solve conundrums in the academic litera-

ture could become an exacerbating factor if the desire to use them proves infeasible. While multi-method research designs are selected, do we have the tools and support needed to successfully carry them out? Based on our collective experiences, room for im-

From a technical perspective, the question is how to handle the high-volume, many-h不做Notes

- Contact information and dissertation summaries for the co-author:
  - Ariel I. Zerai, Graduate Fellow, Center for Democracy and Civil Society, Georgetown University, arielz@georgetown.edu. My dis-
  - Custom, Capture, Collaboration, and Communication: Financial Instruments and Exchange Rate Policies in Eastern Europe, 1990-2004, exam-
  - Jana Cibulkova, Cornell University, jcibulkova@cornell.edu. My dis-
  - On financial system changes. I evaluate this argument using cross-sectional time-series econometric analysis (regression techniques on panel data and duration models) of twenty-five IE countries between 1990 and 2004 and an in-depth examination of the cases of Bulgaria, the Czech Republic, and Poland.
  - Matthew C. Hoffman, University of New Mexico, mmh@unm.edu. My dissertation, Revisions and Extraterritorial Policies: Overex pri-
  - Matt Lieber, Brown University, matt.lieber@brown.edu. My dissertation, Revisions and Extraterritorial Policies: Overex pri-
  - Thomas Pempersky, University of Colorado at Boulder (beginning

non-Western, was underestimated, as was the time it would take to translate our findings into high-volume, many-h

Given our shared understandings, finding effective ways to combine

the essence of being lost at sea when doing our field research. We would have been better served to reflect here on a collection

mechanisms of qualitative methods and the possible pitfalls to avoid.

Once completed, the reception of multi-method work ap-

pears to be mixed. Whether to reach the goal of enriching multi-

method research projects reflects a concern about the quality of scholar-

ship: when different methods are combined and are not

sufficiently mastered. As the number of multi-method research projects increases, the challenge for researchers and students of

methodology begins to nail the discipline's seas, we should be more aware of the destination multi-method research is taking us to,

whether we have effective command of the ships taking us there, and whether we even want to journey there in the first place.

Notes

1 Contact information and dissertation summaries for the co-authors:
  - Ariel I. Zerai, Graduate Fellow, Center for Democracy and Civil Society, Georgetown University, arielz@georgetown.edu. My dis-
  - Custom, Capture, Collaboration, and Communication: Financial Instruments and Exchange Rate Policies in Eastern Europe, 1990-2004, exam-
  - Jana Cibulkova, Cornell University, jcibulkova@cornell.edu. My dis-
  - On financial system changes. I evaluate this argument using cross-sectional time-series econometric analysis (regression techniques on panel data and duration models) of twenty-five IE countries between 1990 and 2004 and an in-depth examination of the cases of Bulgaria, the Czech Republic, and Poland.
  - Matthew C. Hoffman, University of New Mexico, mmh@unm.edu. My dissertation, Revisions and Extraterritorial Policies: Overex pri-
  - Matt Lieber, Brown University, matt.lieber@brown.edu. My dissertation, Revisions and Extraterritorial Policies: Overex pri-
  - Thomas Pempersky, University of Colorado at Boulder (beginning

non-Western, was underestimated, as was the time it would take to translate our findings into high-volume, many-h

Given our shared understandings, finding effective ways to combine

mechanisms of qualitative methods and the possible pitfalls to avoid.

Once completed, the reception of multi-method work ap-

pears to be mixed. Whether to reach the goal of enriching multi-

method research projects reflects a concern about the quality of scholar-

ship: when different methods are combined and are not

sufficiently mastered. As the number of multi-method research projects increases, the challenge for researchers and students of

methodology begins to nail the discipline's seas, we should be more aware of the destination multi-method research is taking us to,

whether we have effective command of the ships taking us there, and whether we even want to journey there in the first place.

Notes

1 Contact information and dissertation summaries for the co-authors:
  - Ariel I. Zerai, Graduate Fellow, Center for Democracy and Civil Society, Georgetown University, arielz@georgetown.edu. My dis-
  - Custom, Capture, Collaboration, and Communication: Financial Instruments and Exchange Rate Policies in Eastern Europe, 1990-2004, exam-
  - Jana Cibulkova, Cornell University, jcibulkova@cornell.edu. My dis-
  - On financial system changes. I evaluate this argument using cross-sectional time-series econometric analysis (regression techniques on panel data and duration models) of twenty-five IE countries between 1990 and 2004 and an in-depth examination of the cases of Bulgaria, the Czech Republic, and Poland.
  - Matthew C. Hoffman, University of New Mexico, mmh@unm.edu. My dissertation, Revisions and Extraterritorial Policies: Overex pri-
  - Matt Lieber, Brown University, matt.lieber@brown.edu. My dissertation, Revisions and Extraterritorial Policies: Overex pri-
  - Thomas Pempersky, University of Colorado at Boulder (beginning