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 and politics, and racial and ethnic politics. One of those 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 APSA dissertation awards, in public law, racial and ethnic politics, political philosophy, comparative politics, and federalism and inter-governmental relations. These dissertations have won other recognitions as well. Thirty-six have been published as university press books to date, with several currently in production. Most of these dissertations have primarily or exclusively used non-quantitative methods.

References

Bedolla, Lisa García. 2005. *Fluid Borders: Latino Power, Identity, and Politics in Los Angeles*. Berkeley: University of California Press.

Collier, David, Henry E. Brady, and Jason Seawright. 2004. "Sources of Leverage in Causal Inference: Toward an Alternative View of Methodology." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, ed. Henry E. Brady and David Collier. (New York: Rowman & Littlefield), 229-66.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Norton, Anne. 2004. "Political Science as a Vocation." In *Problems and Methods in the Study of Politics*, ed. Ian Shapiro, Rogers M. Smith, and Tarek E. Masoud. (Cambridge: Cambridge University Press), 67-82.

Smith, Rogers M. 2005. "Of Means and Meaning: The Challenges of Doing Good Political Science." In *Perestroika! The Raucous Rebellion in Political Science*, ed. Kristen Renwick Monroe. (New Haven: Yale University Press), 525-33.

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

Introduction

Andrew Bennett
Georgetown University
bennetta@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, exhorted by their thesis committees to consider alternative approaches, and above all spurred by the apparent success on the job market of candidates whose research proves they are adept at more than one method.

Yet there are many challenges to using more than one method well, and there is a danger that graduate students in particular will set or be held to unrealistic expectations, that researchers will apply several methods poorly rather than doing one well, and that multi-method techniques will be tacked 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 best to combine different methods in the same research design (Lieberman; Gerring and Seawright; Bennett and Braumoeller). 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 to and synthesize general lessons from the practices of researchers doing empirical multi-method work.

This symposium reflects this state of play. With the exception of myself as its editor (!), it 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 either engaged in multi-method research or seeking 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 on the challenges of doing this demanding kind of research. Although I did not ask the contributors, ranging from current graduate students to senior faculty, to focus on issues specific to their subfields, 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, participant observation, experiments, statistical analysis, case studies, formal modeling, simulations, archival analysis, interviews, and others—and not all of them could be included here. Indeed, one 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 here 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 promise 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 limitiations of the others, a process that Thad Dunning calls "triangulating" in his essay. While some of the contributors may have started off with the intention of doing multi-method work for one reason or another, they were for the most part driven to this practice by the desire to understand a substantively important puzzle 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 having brought a variety of methods to bear. Yet the essays also embody a consensus on the challenges of doing multi-method work as well. Such work can be criticized from all corners on 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 relevant languages and archives, the 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 triangulation, as Thad Dunning points out: Although each method can compensate for the limitations of another, mistakes in any one method can also cumulate when methods are applied sequentially and build upon one another. Accordingly, several contributors share Jason Wittenberg's view that mastery of one method is better than mere facility in several.

The bottom line from these practitioners is that despite the considerable challenges and costs involved, multi-method research is well worth doing. One measure of the high level of interest in this work is that my query to newsletter subscribers for an essay from a current or recent graduate student doing multi-method work led to over a dozen responses. Rather than choosing just one, I asked Scott Siegel, 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 their contact information for readers interested in following up on their particular mix of methods. The resulting essay concisely captures the shared experiences and concerns of this key constituency, noting the considerable professional and intellectual benefits of pursuing multi-method work but underscoring as well the tradeoffs in doing so. One challenge here is attaining appropriate training in different methods, especially those not offered at a high level in every department, and the co-authors of this essay underscore 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 (IQMR). 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. Thad 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 usefully inform the building of formal models. Daniel Carpenter emphasizes this point as well, challenging the use of "as if" assumptions in formal models and urging modelers, as an increasing number of them appear to be doing, to inform their modeling from and test their models in qualitative case studies, rather than just using selected cases to illustrate models.

Susanne Lohmann, 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 Organization*, *World Politics*, the *American Journal of Political 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 tenure committees to judge such work appropriately. Prosaic as these problems are, 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 best address. Wittenberg reminds us that the "single country" study can often be disaggregated into multiple case studies by comparisons across sub-units, over time, or across issue areas. He illustrates this point with his own research, in which his "single-country" study became three thousand observations. Wittenberg concludes with an extremely useful summary of seven issues for scholars to consider when contemplating a multi-method project or fieldwork in a developing country.

Finally, Hein Goemans notes the importance of modeling and explaining historically important individual cases even if their underlying mechanisms do not commonly recur. Goemans stresses the importance of being careful with "off-the-shelf" dataset codings, and he urges scholars to try coding themselves five to ten of the cases in any off the shelf dataset they want to use to see how the codings do or do not fit their own concepts and purposes. Goemans 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 extremely important and promising suggestion for creating wikis for datasets so that users can provide feedback on case codings. This would require working through several difficult challenges—for example, if codings frequently change, researchers will need to keep track of the codings operative on the date on which they were accessed and perhaps periodically 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. None-theless, this offers a very promising approach to getting statistical and qualitative researchers to work together on issues of common concern. Hopefully, conferences and workshops directed to this goal can be organized soon. This proposal is of sufficient importance and magnitude that dataset sponsors and field-wide organizations like the NSF and World Bank need to pool their resources and work together to develop suitable protocols for continuously improving dataset codings (for an excellent listing of dozens of datasets on international relations and politics, see http://garnet.acns.fsu.

edu/~phensel/data.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 methodological communities dining at separate tables, but fostering this development requires significant institutional changes to make it easier to carry out and publish multi-method work, 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 unrealistic expectations, especially for our 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 accommodate the higher word counts and/or Web-based appendices that such work can require. Departments need to ensure methodological pluralism in hiring and promotion, and to accommodate the fact that some research agendas are better suited to articles and others to books. Departments also need to ensure that their graduate students have access to and resources for cross-method training, either in-house or through dedicated methods training programs. These programs, in turn, need to incorporate multi-method approaches not just into their curricula but 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 collaborations among scholars with diverse kinds of methodological and substantive expertise. Like the pain of childbirth, the hardships endured in undertaking multi-method research appear to recede in memory as time goes on and the benefits become self-evident. It just doesn't have to be so painful in the first place.

Title Here

Hein E. Goemans
University of Rochester
hgoemans@mail.rochester.edu

In this brief essay I will elaborate on some of the points I raised at last year's APSA panel on multi-method work. As always, I emphasize the essential complementarity of different methods. I first briefly discuss why qualitative research and formal models have much to offer each other and why scholars in each methodological tradition can gain much from a better understanding of the other tradition. I then shift to a focus on the overlooked link between qualitative and quantitative research, to argue for a reconsideration of the requirements of the inputs of quantitative research: What constitutes good data? I close with a proposal that calls for collaboration between qualitative and quantitative researchers to set new standards for the collection and use of large-N datasets.

I propose that qualitative research is a natural partner of formal models and *vice versa* because case studies can actually track the mechanism proposed by the formal model. Until recently, scholars preferred to test the predictions of their formal models with the usual statistical methods. However, as Signorino (1999) showed, the usual statistical methods are very poorly suited to test formal models because they assume observations to be strategically independent (conditional on the explanatory variables). The strategic formal models in political science, in sharp contrast, draw their strength from the insight that decisions are strategically *inter*dependent. It is clearly inappropriate to test a theory which poses strategic interdependence of observations with a method that assumes observations are independent. Signorino (1999) and Lewis and Schultz (2003) offer a way out of this predicament by showing how to develop fully structural estimators designed to address the issue. This approach is also advocated in the influential EITM (Empirical Implications of Theoretical Models) workshops. This approach, however, has two distinct drawbacks. First, it assumes that the formal model represents the "Truth" 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 results—even if run on the same data. This is an extremely heavy load for a model to bear—as most modelers would admit. Second, it assumes the model represents a pattern that regularly appears. However, the strength of a formal model does not derive from its ability to explain a great many cases, as long as it can explain some (hopefully important) cases that other models can not explain. Formal models of war, for example, do not necessarily claim that all wars are caused by their particular mechanism, only that the mechanism occurred at least once and could occur again.

In comparison to both the older and more recent statistical methods, the case study method seems 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 Schultz (2001) and Goemans (2000), case studies can trace not only which choices were considered and actions were taken, they can also show that some other actions were deliberately avoided in anticipation of the choices and actions of the other player(s). Moreover, 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 inherent complexity of the real world and trace specific causal mechanisms. Case studies can trace and establish causal mechanisms in the midst of a potentially overwhelming number of otherwise confounding factors. Even if the empirical process 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 generality. To show the relevance and power of a model, it suffices to empirically trace the causal mechanism in a handful of cases; even one—preferably "important"—case can do. A statistical search for a general, statistically 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 suffice to demonstrate the usefulness of a particular model with only a few cases. Moreover, some empirical patterns are (extremely) rare. Some rare events, however, are extremely important and deserve both theoretical and empirical attention. If there are only a handful of cases of a particular kind of important 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 decisions. This does not mean, however, that such a theory is unimportant or undesirable. I find it entirely plausible that formal models 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 partners for formal models, they are *essential* partners for quantitative research. Graduate students and faculty make enormous efforts to acquire the most up-to-date, most advanced toolbox of statistical methods. Prominent journals publish relatively narrow articles that deal with some (minor) perceived statistical problem and its potential solutions. Without question, these articles can be useful and extremely helpful. However, this sophisticated technological arms race has gone on with little or no attention to the inputs: the actual data to be analyzed. Scholars take data "off-the-shelf" with little thought or consideration of the purposes for which the data were originally collected, the coding schemes and decisions, and the reliability and accuracy of the data.

For example, the Correlates of War data, as well as the International Crisis Behavior and Militarized Interstate Disputes data on conflict, record the outcome of conflict, and all three datasets in essence record the *military* outcome of conflict. However, for researchers who adopt the bargaining model of war, the military outcome of a conflict may be (almost) irrelevant; instead such scholars want to know whether at the end of the conflict the leader or state was better off than before.[1] Thus, the 1973 Yom Kippur War may count as a military defeat for the Egyptians, but the Egyptian crossing of the Suez Canal and overall military performance came as a shock to the Israelis. 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 full documentation for coding decisions then can make it difficult for other scholars to understand the data at any significant level of detail. We are simply to trust that a group of (often) undergraduate students correctly implemented the coding rules, although they may have little or no understanding of the relevant concepts and purposes of the data. Nevertheless, off-the-shelf data is almost always taken at face value, as capturing and measuring the essential concepts of interest and a complete sample. It is not only at their own but also at their reader's peril that quantitative scholars ignore that each data point entry in essence constitutes a mini case study, which requires full documentation. In essence, the construction of a data set requires *qualitative* skills.

While relatively few scholars attempt to construct their own data sets from the ground up, the discipline would significantly benefit if not only creators but also *users* are held to a higher *qualitative* standard. For my own students I have instituted the following rule: If you rely on off-the-shelf data, take between five and ten random observations[2] from the dataset and attempt to code the variables *from the ground up*. For datasets that have only one variable per observation—for example, data that only record the existence of an armed conflict—students are required to pick a time span within the domain of the original data and five randomly chosen countries and code the presence or absence of an armed conflict *from the ground up*. Students are required to provide full documentation of their codings. This gives students a good first-hand idea of what went into the data, as well as how appropriate and accurate the data are for their own research. At the same time, it requires students to do at least a minimum of historical and qualitative research.

The discipline as a whole would probably benefit if we were to require such a protocol from our students for all quantitative research. In order to harness such efforts, it might be a good idea to collect, organize, and integrate information on dataset codings thus generated by a community of scholars in a systematic fashion on a wikisite. Just as wikis like Wikipedia invite distributed users to provide input that editors can vet and then integrate into evolving online documents, data set wikis could incorporate input from widely dispersed experts to improve and qualify case codings. Although Wikipedia has at times suffered from the short-term appearance of fraudulent information, a study in the journal *Nature* found it to be nearly as accurate as the Encyclopedia Britannica (Johnson 2006). A recent study of the accuracy of datasets on democracy in Central America indicates the potential value of incorporating input from diverse experts into case codings (Bowman, Lehoucq, and Mahoney 2005).

Qualitative as well as quantitative researchers would be invited to contribute to this effort. Qualitative research on some cases often produces information on variables of interest in quantitative data sets, and could thus provide invaluable documentation and background information on some the codings of various datasets. Over time, and with help from students elsewhere, the wiki would "backwards engineer" and fully document the most important and most often used datasets. Such a collective effort would provide a much-needed and extremely rich qualitative backbone to quantitative datasets.

Notes

[1] Full disclosure: In the absence of a better alternative, I have used the military outcome of conflict as a proxy for the political bargaining outcome myself.

[2] The number of observations depends on the number of variables to be coded for each observation.

References

Bowman, Kirk, Fabrice Lehoucq, and James Mahoney. 2005. "Measuring Political Democracy: Case Expertise, Data Adequacy, and Central America." *Comparative Political Studies* 38:8 (October), 939-70.

Goemans, Hein E. 2000. *War and Punishment; The Causes of War Termination and The First World War*. Princeton: Princeton University Press.

Johnson, George. 2006. "The Nitpicking of the Masses versus the Authority of the Experts." *New York Times* (January 3), F2.

Lewis, Jeffrey B. and Kenneth A. Schultz. 2003. "Revealing Preferences: Empirical Estimation of a Crisis Bargaining Game with Incomplete Information." *Political Analysis* 11:4 (Fall), 345-67.

Schultz, Kenneth A. 2001. *Democracy and Coercive Diplomacy*. New York: Cambridge University Press.

Signorino, Curtis S. 1999. "Strategic Interaction and the Statistical Analysis of International Conflict." *American Political Science Review* 93:2 (June), 279-97.

The Trouble with Multi-Methodism

Susanne Lohmann
University of California, Los Angeles
lohmann@ucla.edu

Quick, pick the statement that best represents practice in political science: (1) To the political scientist who comes equipped with a hammer, everything in the world presents itself as a nail; (2) the political scientist employs multiple methods, tailoring the mix to illuminate from various angles the phenomenon to be explained.

The first statement rings true; not so the second statement. Political science favors the deeply specialized over the multi-methodist. The specialized political scientist knows everything about nothing and nothing about everything. He regularly hangs out with similarly specialized political scientists. They are his kin. When he submits a paper to a journal or applies for a grant or comes up for tenure, he can rely upon his "friends and family" for a favorable review. Meanwhile, the political scientist who takes a multi-method approach knows a little about a lot or a lot about a little (depending on whom you ask). She lives on the fringes of several specialized clusters. She connects the clusters to each other, but is herself only loosely connected to any given cluster. When she submits a paper to a journal or applies for a grant or comes up for tenure, her specialized reviewers compare her paper submission or her grant proposal or her tenure record to the papers or proposals or records of their specialized kin, and she inevitably comes up short.

All else equal, the multi-methodist political scientist will end up with fewer publications, fewer publications in leading political science journals, fewer grants, and weaker tenure letters—a weaker tenure case, that is, compared to the deeply specialized political scientist. It is thus that the multi-methodists die and the deeply specialized multiply.

This is a problem, for it takes a mix of both types—the deeply specialized and the multi-methodist—for the scientific process to enjoy vibrancy and for scientific progress to occur. This essay spells out the trouble with multi-methodism and explains why there is no obvious structural solution to the problem—all we can do, really, is educate journal editors and foundation officers and department chairs and academic deans.

Why *Fachidioten*?

The concept of a *Fachidiot* emerged in tandem with the 19th century German research university. This German expression translates literally as "idiot of his department" and more freely as "he who knows everything about his little department of knowledge and nothing about the whole wide world."

The fundamental force driving *Fachidiotentum* in the research university is the fact that the whole wide world is too big to fit into any one little brain, and so the world must be parceled up and distributed across lots of little brains. Thus, one scientist specializes on this kind of rock, another, on that kind of ant, a third, on yonder kind of legislature, and hopefully all of their scattered insights will add up to a coherent picture of the world.

The function of the research university is to enable deep specialization. The protectives structures of the university evolved over the centuries to solve several problems: how to nurse deeply specialized scholars, how to protect them from each other and the outside world, and how to pool the results of their distributed inquiries (Lohmann 2003, 2007). The university as it currently stands does a better job protecting its *Fachidioten* than it does pooling their partial insights.

Inevitable as it is, *Fachidiotentum* comes with two important downsides. First, deeply specialized scientists tend to be reductionists. They do an excellent job of analyzing the individual components of complex phenomena in isolation from one another, and they do a horrible job of "putting it together."

Second, *Fachidioten* tend to get stuck in methodological ruts. Precisely because they hang around with like-minded specialists, they tend to stick to their guns (or hammers) even when the method in question has exhausted its potential.

There exists no alternative to *Fachidiotentum*, but a complementary force is available. What multi-methodists have to offer, first, is a holistic picture of the overall functioning of the phenomenon in question. Second, as multi-methodists hop from one specialized cluster to another, they inseminate any one cluster with the ideas they picked up in the other clusters, which is what drives methodological renewal in all clusters.

My argument holds true, too, for methodological approaches in political science: formal modeling, regression analysis, survey research, laboratory experiments, computer simulations, close reading of texts, case studies, and historical analysis—so many methods, so little time. Thus, political scientists end up specializing in one kind of method, and hopefully all the various methods will come together in the end—except, of course, that they don't. For this reason, there is an urgent need for a mix of multi-methodism and deep specialization.

Incentive and Selection Effects

This is where we run into a disconnect between what's good for political science and how political science works. Political science thrives when it can rely on a mix of deep specialization and multi-methodism: the ideal is not either-or, but both. And yet there are powerful incentive and selection effects in political science that favor 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 academic 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 presses. In practice, the relative importance of articles and books varies across subfields and sub-subfields of political science. American Politics relies 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 illuminate a phenomenon in full. The effect I describe below is thus more powerful for articles and weaker for books. 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 heavily 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 medicine. Even so, they count for a little bit. Once again, we find variation within political science: the subfields and sub-subfields that are fond of articles tend to look favorably upon National Science Foundation grants—which is not a coincidence, as we shall see, for the types of political scientists who find it (relatively) easy to publish articles also have a (relatively) easy time getting National Science Foundation grants.

Let us take a look at publications in leading political science journals. Suppose you submit a deeply specialized paper to a leading journal. The editor will select a handful of reviewers who are steeped in the same specialization, meaning they share the methodological framework that underpins your paper. The reviewers might still shoot down your paper (after all, it can be a bad paper). But they won't reject the paper on fundamental methodological grounds. Instead, they will fuss around at the margins of your method, and you can easily address their little concerns.

Now suppose your paper employs a multi-method approach: you mix and match—say—game theory, a historical case study, and a regression analysis. The editor will assign a motley crew of reviewers, consisting of one game theorist, 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 come up short. They might still recommend your 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 so you will get bogged down in an endless sequence of revise-and-resubmits where 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 could respectfully accept the fact that other political scientists use different methods, or an eclectic mix of methods. In practice, political scientists tend to conceptualize political scientists who use different methods as *the wicked other*. Thus, a specialized reviewer who is charged with evaluating a multi-method paper will tend to project hostility rather than indifference, let alone support.

The journal editor might understand the workings of the bias against multi-method papers and undo the bias by weighing referee reports on multi-method papers differently than she does referee reports on specialized papers. Then again, she might not, in which case she will end up accepting more specialized papers than multi-method papers.

In practice, some journals support deeply specialized scholarship (*Journal of Conflict Resolution*), others, multi-methodist scholarship (*World Politics*). The flagship disciplinary journal can go either way, and in the recent past the *American Political Science Review* has chosen to go with the deeply specialized. Fights over the focus of the *American Political Science Review* tend to reflect a cleavage between quantitative and qualitative methods (e.g., rational choice and regression analysis versus case studies and historical analysis). But the quantitative-qualitative cleavage captures only part of the action; another important cleavage lies between the deeply specialized and the multi-methodist. Quantitative methods and deep specialization 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 publications, 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 revise 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, the multi-methodist political scientist will be lacking, 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, to which I turn now. 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 for a multi-methodist will be shorter and all over the place, for 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-methods candidate for tenure—whom they don't know all that well in the first place—to "their own," with the effect that the multi-methodist once again comes up short.

The faculty in the candidate's academic department who vote on the candidate's tenure case might 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 with deferring to the deeply specialized faculty who are 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 scholarship, they like to count the number of publications weighted by journal, the number and size of grants, and the number of favorable 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 themselves too thin." In practice, it seems that people's types are relatively fixed—they are who they are and do what they do, and they either survive or die. This is why the bias in the tenure process tends to create a selection effect rather than an incentive effect. Post-tenure, the deeply specialized political scientists will dominate the landscape at the expense of multi-methodists, with detrimental consequences for the vibrancy of political science.

Given the structural bias against multi-methodists, the average multi-method political scientist who makes tenure at a leading research university will be better than the average specialized political scientist. It immediately follows that multi-methodists will be disproportionately represented among the leading intellectual figures in political science. Young political scientists who emulate "who's out there doing the best work" can thus go seriously astray (if you want to call it that). From the fact that some of the best political scientists are multi-method scholars, young political scientists might infer that multi-method scholarship is valued in political science, which is certainly not true for the tenure process, and they might make faulty career decisions based on this incorrect inference.

Enter the Fundamental Attribution Error

One problem that besets inferences about scientific review processes is the fundamental attribution of social psychology, which describes the human propensity to explain the behavior of other people with reference to people's characteristics as opposed to the characteristics of the situations people are embedded in, as in, "she failed because she is incompetent" as opposed to "she failed because the process was biased against her." The fundamental attribution error is reversed when people reason about themselves: If I fail, it is because the process was biased against me, and not because I am incompetent. Interestingly, women are less likely to fall for this reverse error; when they fail they are more likely than men to attribute their failure to failings of their own rather than failings of the system they are embedded in, which arguably includes biases against women in science.

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, because I have published both single-method articles (e.g., Lohmann 1992, 1993, 1994b, 1997, 1998a) and multi-method articles (e.g., Lohmann 1994a, 1994c, 1998b), and so I can directly 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 in my mind that it is harder by an order of magnitude to publish a multi-methods paper than 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 such as Ford and Templeton, which care about making a difference 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 political science discipline—National Science Foundation grants—are elusive.

Third, all of my promotions have been difficult, not primarily because of my multi-methods scholarship (I have enough single-methods research to show), but because my single-methods scholarship cuts across two disciplines (political science and economics), and it turns out that my argument about the bias against multi-method political scientists holds 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 typically takes for the university to process my personnel reviews (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 scholarly or creative achievement. My department promptly suffered a nervous breakdown and ended up voting to recuse itself from my personnel review, as a result of which political scientists 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, however baffled and battered, and one of the wonderful things about tenure is, of course, that once you have it, you are amazingly free to do whatever research you consider valuable.

To the *Fachidioten* with his ahistorical mindset, the protective powers of tenure are hidden. Case in point is Steven Levitt, an economics professor 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 of Chicago to revoke his tenure in exchange for a $15,000 salary increase. His best-selling book *Freakonomics* is subtitled "a rogue economist explores the hidden side of everything" (Levitt and Dubner 2005), but Levitt would hardly have been hired and promoted as an economics professor 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 dismally conventional, which is presumably why he is so very willing to give up tenure (whether he would really do so if Chicago took him up on his offer is another question). Multi-methods faculty like myself, who have seen the dark side of academic departments, are deeply appreciative of tenure.

One question in my mind is whether my troubled promotions at the Los Angeles campus of the University of California are due to my gender: to this day, many men—and women, too—have problems accepting excellence when it comes in female form. In my case, issues of interdisciplinarity and gender are impossibly entangled. (Yet another question in my mind is whether gender correlates with interdisciplinarity, or multi-methodism, due to sex differences in cognitive functioning, as in, the male brain is autistic, the female brain, multi-dimensional.) Even if in my case no clean inference is possible, there exists quite a bit of evidence that scientific peer review processes work to the disadvantage of women.

Case in point is a 1997 *Nature* article titled "Nepotism and Sexism in Peer-Review," which was co-authored by microbiologist Christine Wennerås and immunologist Agnes Wold at Göteborg University in Sweden. Two years earlier, Wennerås and Wold had applied to the Swedish Medical Research Council for research fellowships and had been rejected. In the 1995 round, 52 of 114 applications were female, but only four of the 20 fellowships were awarded to females. Wennerås and Wold had to go to court to get access to the peer reviewer's evaluations of the applicants on three dimensions: competence, relevance of the proposal, and quality of the experiment. They compared the applicants' scores with the applicants' publication records, taking into account both the number of articles published and the prestige of the journals in which the articles were published. Wennerås and Wold found that reviewers consistently gave women lower scores, especially on competence. In this set, a women applicant needed three additional publications in journals like *Science* or *Nature* to be scored equally competent with a man. This stunning result triggered a debate among European science leaders. Today, we are seeing the beginnings of a paradigm shift away from blaming individual women and towards understanding bias in the scientific construction of excellence and the exclusionary mechanisms of scientific elites.

So this is the message I would like to get across to political scientists who are just starting out on the long trek to tenure. Multi-methodist political science is valuable, but you must know that, pre-tenure, the deck is stacked against you, and you must understand the incentive and selection structures to deal with them intelligently. Especially the women among you need to be careful not to blame themselves when they meet with external obstacles.

It's Better to Rely on Well-Designed Institutions than on Well-Behaved People

One characteristic of a political scientist is the desire to come up with structural solutions to problems. One solution to problem of multi-methodism would be for multi-methodists to cluster together and review each other's paper submissions and grant proposals and tenure records. There are two problems with this solution.

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 downsides of *Fachidiotentum*. 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, over time they would become detached from 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 connectedness to the deeply specialized political scientists, and it shows.

There is no obvious structural solution to the problem of the bias against multi-methods 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 wimpy solution, admittedly, but there it is.

References

Levitt, Steven D. 2007. "Let's Just Get Rid of Tenure (Including Mine)." Freakonomics Blog. http://www.freakonomics.com/blog/2007/03/03/lets-just-get-rid-of-tenure/. Accessed 4-19-07.

Levitt, Steven D., and Stephen J. Dubner. 2005. *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything*. New York: William Morrow.

Lohmann, Susanne. 1992. "Optimal Commitment in Monetary Policy: Credibility versus Flexibility." *American Economic Review* 82: 273-86.

Lohmann, Susanne. 1993. "A Signaling Model of Informative and Manipulative Political Action." *American Political Science Review* 88: 319-33.

Lohmann, Susanne. 1994a. "Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989-1991." *World Politics* 47: 42-101.

Lohmann, Susanne. 1994b. “Information Aggregation Through Costly Political Action.” *American Economic Review* 84: 518-30.
Lohmann, Susanne, and Sharyn O’Halloran. 1994c. “Divided Government and U.S. Trade Policy.” *International Organization* 49: 595-632.
Lohmann, Susanne. 1997. “Linkage Politics.” *Journal of Conflict Resolution* 41: 38-67.
Lohmann, Susanne. 1998a. “An Information Rationale for the Power of Special Interests.” *American Political Science Review* 92: 809-27.
Lohmann, Susanne. 1998b. “Federalism and Central Bank Independence: The Politics of German Monetary Policy, 1957-1992.” *World Politics* 50: 401-46.
Lohmann, Susanne. 1999. “What Price Accountability? The Lucas Island Model and the Politics of Monetary Policy.” *American Journal of Political Science* 43: 396-430.
Lohmann, Susanne. 2003. “Darwinian Medicine for the University.” In *Governing Academia*, Ronald G. Ehrenberg, ed. (Ithaca, NY: Cornell University Press), 71-90.
Lohmann, Susanne. 2007. *How Universities Think: The Hidden Work of a Complex Institution*. Manuscript. University of California, Los Angeles.
Wennerås, Christine, and Agnes Wold. 1997. “Nepotism and Sexism in Peer-Review.” *Nature* 387: 341-3.

Some Observations on the Interweaving of Model and Narrative

Daniel Carpenter
Harvard University
dcarpenter@gov.harvard.edu

Model-driven political science has had some stunning triumphs in the last few decades. Whether the model is verbally expressed (as in Theda Skocpol’s theory of states and social revolutions or David Mayhew’s theory of reelection minded legislators) or mathematically formalized, theoretically-driven work has added richly to the clarity of many political science studies and has allowed for aggregation of findings and studies to help form a more coherent and comprehensive body of knowledge. While I prefer a combination of historical and theoretical research to historical inquiry alone, I often find it helpful to approach research questions by viewing them partially (never wholly) through a formal lens, often mathematically.[1]

There is a rub, though. If there is a weakness in model-driven political science, and there are many, it lies in the debility of its “causal mechanics.” To demonstrate this, I select a reasonably common event in the contemporary world: A politician gets caught in an extramarital affair or some form of sexual or romantic indiscretion. Or a statement or action of the politician is exposed as dishonest. The event is publicized. Voters (at least in the United States) turn against the politician, and support for his or her party drops precipitously. At the end of the day, our scandalous politician loses office (or his or her allies do).

One common reading from the theoretical literature for patterns such as this–probably the most common reading–is that they represent “retrospective voting.” Some unobserved aspect of the candidate’s “quality” has been revealed to voters, and now that they are better informed they revise (in some sort of Bayesian fashion) their beliefs about him, and with those revisions headed downward, they are less likely to vote for him in whatever optimum or equilibrium obtains.[2]

Yet when we think about the *mechanics* of how we got from “A” (scandalous politician) to “B” (angry voters), then a retrospective voting account of this pattern is deeply unsatisfying. The idea of causal mechanics is not new, but one might think of a causal graph with a density of intermediate nodes that correspond to different “mechanisms” that must be operative in the process described (perception of a politician’s action, emotional and rational response, mental deliberation or conversation with other agents about the action, belief/opinion formation, determination to act, action).[3] What actually goes through agents’ heads when they enter the voting booth to vote against the lying, cheating politician, or when they participate in some meaningful way against him? If rationality drives them, do they actually compute utilities (as many of our theorists portray them doing) and compare possible strategies according a maximization (or “best response”) criterion? Do they truly use Bayes’ rule and conditional probability to update? Or is the data in a voter’s head much more symbolic and categorical? If voters’ judgments and attitudes about policies and politicians can be expressed and measured using “thermometers,” then what is the mechanical role of emotions in the revision of beliefs about a politician?

Our inability to speak clearly to questions such as these stands as evidence that political science lacks the mechanically compelling portraits of behavior that animate psychology, cognitive science, and neuroscience. Part of the problem lies in our practice of taking problematic concepts and states of the world and slapping labels upon them (which means we have not explained them at all). We can describe the wronged citizen as “loss-averse” or “risk-averse,” but that is an act of description convenient only for simplifying a model.

Another response to worries such as this is the old argument that the postulated behavior of a model is “as if” behavior. As long as the voter behaves as if she is punishing the politician, or as long as she behaves as if she is updating her beliefs about him in a Bayesian fashion, then all is well with our explanatory paradigm. The extreme form of this lies in Milton Friedman’s now tired argument that models and theories should be judged not at all on the accuracy of their assumptions but on the accuracy of their predictions. Despite the allure, this is in reality a fundamentally unscientific way of approaching political reality.

If we admit that questions of causal mechanics such as this are compelling and need to be addressed fully, then it will soon become clear that much (not all) quantitative research is not of much help. At least as it is currently structured, examining aggregate voting behavior in the wake of a scandal is likely to tell us little about what goes on in inside the brain. It will tell us next to nothing about the involvement of emotion. Neuroimaging research is one promising avenue with which such questions can be addressed, but even then the success will be only partial. For one, brain imaging suffers from massive challenges to external validity–it is difficult if not impossible to create or recreate valid and portable conditions of politics in the neuroimaging laboratory. For another, and perhaps more important, the sorts of inferences made in neuroimaging place the analysts at several removes from inferences on the order of "this part of the brain lights up, we conclude X, that part of the brain lights up, we conclude Y," as one neuroimaging specialist explained it to me.[4]

As heretical as it might sound, I want to suggest that historical narrative–of the sort practiced by careful historians and (fewer) careful and historically-oriented political scientists–can assist us in tasks such as these. Suppose I attempt to study the "scandalous politician–angry voter" pattern not by running an experiment or by examining aggregate vote choice after a particular scandal, but by carefully and broadly analyzing speech and action in the wake of a scandal such as Watergate or Teapot Dome.

What sort of data will such an inquiry produce? It might produce lots of images and symbols that the historical analyst sees scattered around the culture–a negative advertisement, a surly photograph, a tearful apology on television, a defiant denial in print. It might produce lots of text–editorials, letters, Web-log entries, diaries, news stories, police logs, administrative records–from which I can recover some of the likely content of voters' thoughts, emotions, and actions. These sorts of "data" have all sorts of disadvantages–they are difficult to compare in an easy statistical manner, speech might misrepresent "true" feelings and intentions, and texts might be subject to multiple interpretations–but they have advantages that many of our quantitative data do not. For one, we can see some part of the content of voters' thoughts by looking at the media images or texts ("I did not have sex with that woman, Miss Lewinsky"; "Read my lips, no new taxes") that they were almost certainly processing. For another, we can often observe emotions and actions taking shape. We can see a person reasoning through different courses of action, we can observe (some of) the conversations and deliberations in which new attitudes, opinions and preferences are formed.

To do this, of course, requires that we recognize the ability to extract information from symbolic acts and speech acts. Much of modern political science seems founded upon a skepticism of this sort of move, in part because of underlying concern that "talk is cheap" and preferences are revealed only through observed costly action. Yet historians and ethnographers have long stood in recognition of this point and have nonetheless rendered fascinating contributions from analysis of texts (diaries) and symbolic complexes (advertisements, works of pictorial art). The point is this: *Even if talk is intentionally misrepresentative, it may be informative*. I was introduced to this possibility in reading Harvard historian Walter Johnson's wonderful narrative of the New Orleans slave market in the 1840s and 1850s (*Soul by Soul: Life Inside the Antebellum Slave Market*). One of the most enlightening passages was Johnson's shrewd elaboration of a sort of methodology for reading lies (*Soul by Soul*, 12). One of Johnson's sources is a set of court records (regarding contested slave sales) that have been only recently discovered. Yet Johnson recognizes that he cannot validly take these texts at face value. Instead of discarding them (as would most social scientists and historians), he instead exploits the subtle and embedded information they contain.

> ...I have generally read the docket records as if they contain only lies. And yet lies, especially sworn lies given in support of high-stakes legal action, must be believable in order to be worth telling: these lies describe the circumstances of a specific sale in the terms of a shared account of what was likely to happen in the slave market.

Johnson follows this point with a rehearsal of the "few stock stories" that were repeated time and again in court cases. When traders, owners, and slaves themselves would lie about slaves' bodies–a slave overstating his weight or understating his age, an owner assigning virginity to a woman who would be prized by a buyer for her child-producing potential (or for sexual domination by the owner)–these lies persisted but tell us how the process of "commodification" worked. The slave market, as Johnson shows, melded the most perfected mechanisms of the capitalist world (the price system, the meeting of buyer and seller in a differentiated market) with the power system of race and domination in Southern society. It converted "people into prices." And the evidence that commodification occurs comes not from a quantitative examination of price, but *from repeated and competing lies told about slaves in court battles*. The idea that a lie has some informational value for the reader/historian, value that can systematically be appraised, even with error, was very novel to me. I'm currently writing a book, and in confronting the ambiguous statements that pharmaceutical companies, clinical pharmacologists, and FDA regulators told each other over the past half-century, Johnson's book–topically a light year's remove from my project–made me rethink what I was reading.

There are, finally, patterns of human activity and meaning that most of our existing models are poorly equipped to make sense of. Two of these include (1) the persistence of misunderstanding across individuals and across cultural groupings, and (2) learning from rare and symbolic events.

A quick look at our own world should convince us of the relevance of these stable misunderstandings, but driven by equilibrium analysis, rational choice models are generally poorly equipped to handle these realities. How do we come to terms with cultural and political misunderstanding–not just occasional errors of meaning but persistent, decades-long (even centuries-long) cultural orthogonality of the sort that prevailed in the Great Lakes region of colonial North America from the mid-1600s to the early days of the American Republic? This durable moment in time-space was the venue for the meeting of dozens of rival and distinct cultures. Some of these cultures were composed by radically different Native American peoples (the Iroquois and the various nationalities allied against the so-called "Algonquians" such as the Fox, Kickapoo, Ojibwa [Chippewa], and Miami. At least two other cultures were the peculiar settler and imperial cultures that were spawned by the French and British presence in colonial North America. In a wonderful book by Stanford historian Richard White–*The Middle Ground: Indians, Empires and Republicans in the Great Lakes Region, 1650-1815*–the limits of our existing models for explaining centuries-long realities become clear. In a land where no organized entity (no Indian nation, no Western imperial state outpost) was in complete control, and where dependence upon the harsh environment was great for all concerned, the world was governed by a politics of mutual deference and, importantly, consistent misunderstanding. These cultural realities did not merely "affect" the fur trade and the French-Algonquian alliance–they constituted it. The language of alliance was one of "artful manipulation" (*The Middle Ground*, 152), and the alliances were held together by symbols (the calumet, or the atonement ritual) that were interpreted quite differently by different facets of the alliance.

Or consider how those who design institutions learn about them. How did eighteenth-century politicians learn from the experience of state constitutions with legislative supremacy–as narrated in Gordon Wood's magisterial *Creation of the American Republic*–and adjust their beliefs to conclude that strong executive power was necessary in a mixed regime? In some respects the historical experience under state constitutions in the years 1776-1780 served as "raw data" for later founders such as John Adams, Alexander Hamilton, and James Madison. In other respects the data came from "philosophy"–a reconsideration of the celebrated arguments in favor of executive power in a mixed regime, from Baron de Montesquieu's *Spirit of the Laws*. Yet these "data" were never available numerically to actors of the time and were never "learned" using rules of conditional probability. The evidence base consisted instead of rare events, sometimes single events, that were "observable" only symbolically, and even then were subject to various interpretations. Wood's triumph is to draw out a pattern of learning through observation, deliberation, and political conflict, a pattern constructed from a wealth of primary source materials such as pamphlets, editorials, essays, petitions, newspapers, and broadsides.

My sum point is that primary-source-driven narrative and mathematical modeling can complement one another in ways that render the status quo–modeling combined only with statistical estimation alone–a highly impoverished research agenda.[5] In order to realize the possibilities for weaving model and narrative together, practitioners of each single art will need to recognize the limits of their own approaches and how they can be complemented by more imaginative research practices.

Notes

[1] Key for me is that mathematical modeling need not be modeling that is committed in any way to the rationalist paradigm. Some of the most fascinating mathematical models are those that explore "bounded rationality," network dynamics, stochastic processes, or something else. The paucity of such non-rational formal models in political science is a material weakness of our discipline.

[2] Alternatively, their vote against the candidate or his party may comprise some sort of "trigger strategy" or generalized form of punishment for the politician's deviation from cooperative or truthful behavior. For an empirical analysis of one such scandal which shows that retrospective voting is far more complicated than our current models would suggest, see Michael A. Dimock and Gary C. Jacobson, "Checks and Choices: The House Bank Scandal's Impact on Voters in 1992," *The Journal of Politics* 57:4 (November 1995), 1143-59.

[3] For discussion of graph-theoretic approaches to causality, see Judea Pearl, "Causal Diagrams for Empirical Research," *Biometrika* 82 (1995), 669–710 (with discussion); Pearl, *Causality: Models, Reasoning and Inference* (New York: Cambridge University Press, 2000); Steffen L. Lauritzen, *Graphical Models* (Oxford: Clarendon Press, 1996).

[4] For a wonderful account of how the apparatus of experimentation can place scientists at further and further remove from the material of their inquiry, see Peter Galison, *Image and Logic: A Material Culture of Microphysics* (Chicago: University of Chicago Press, 1997).

[5] The most direct attempt to meld narrative and rational-choice modeling was the *Analytic Narratives* volume by Bates, Greif, Levi, Rosenthal, and Weingast. The unfortunate feature of that effort was its absence of narrative–too many of the narratives were in fact data analyses, and it was difficult to separate what the authors called "analytic narrative" from "theory generates comparative statics when the meet data" exercises with which we are very familiar in modern political science.

Peril and Promise: Multi-Method Research in Practice

Jason Wittenberg
University of California, Berkeley
witty@berkeley.edu

There's no doubt about it: multi-method research is in vogue.[1] Perhaps the most obvious evidence of this comes from the job market. Job candidates who successfully combine multiple approaches get that ineffable "buzz" and are often showered with adulation and, ultimately, job offers. As one faculty friend opined with regard to one particularly exceptional candidate, "she is clearly a new kind of comparativist." Another remarked that the work was so good, "the talk could have been delivered in Greek." Graduate students have taken these signals to heart. Increasing numbers are attempting to master qualitative, quantitative, and formal approaches and to formulate methodologically eclectic research proposals. Yet it bears noting that the road to the Promised Land isn't so easy to navigate. Creatively and effectively combining multiple methods is time-consuming and difficult, as the article in this newsletter by current and recent PhD students doing multi-method research attests (Siegal et al., 2007). It can also be risky. In my experience, search committees prefer candidates with mastery of one method to those with mere facility in 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 dissertation and ultimately my book, *Crucibles of Political Loyalty: Church Institutions and Electoral Continuity in Hungary* (Cambridge University Press, 2006). The target audience is graduate students and others who want to employ multiple methods but find the prospect 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 mid-course improvements as circumstances demand. 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 Latin America. This discovery opened up 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 sundry 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 elucidated 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 heard it here first!). Comparative politics was regionally subdivided, and my tacit assumption was that cross-regional comparisons, while technically possible, were of limited analytic utility. Latin American and Southern European authoritarianisms seemed too different from East 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 impractical and ultimately of 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 quite enough. I knew early on that this choice might pose problems of generalizability, but I came to realize that the only way to address the larger puzzle of long-term political continuity would be to explain how it played out in one 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 quantitative and qualitative analysis, I would have to say that I 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, but the focus was on the research question rather than methodological eclecticism per se. That I nonetheless ended up employing an array of tools is evidence that then-existing folk wisdom and common sense on 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 using multiple methods, at least in some form, long before they achieved their current exalted status.

Why should old patterns of mass political loyalties 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 objection. The first and scariest (at the time) was the charge that Hungary was not an interesting place to explore the question. The "why did you study <country-name>" question is among the most common one encounters, and woe unto whoever cannot provide a satisfactory answer. My response was to situate Hungary as a "least likely" case to exhibit political continuity. Prior theory tended to focus on the less disruptive authoritarianisms of Latin America or Southern Europe, where the covert activities of parties, trade unions, and other organizations opposed to the dictatorship were invoked to account for partisan persistence. Under communism 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) refer to as the "n=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 results suspect. A lower level of aggregation was required. I spent a good chunk of my first extended period of fieldwork attempting to gather such data. Ultimately I succeeded in collecting and matching pre- and post-communist municipality-level electoral data for the entire country. One "case" yielded nearly three thousand observations. Through basic quantitative analysis I established many patterns of political continuity and discontinuity.

Finally, the price of eschewing cross-national breadth had to be paid in explanatory depth. The comparative statics were novel and fascinating, but for social science the real value-added of the project was in illuminating how pre-communist partisan attachments were successfully transmitted into the post-communist period. This required returning to Hungary in the hope that I could find materials that would shed light on the differences between localities where there was partisan persistence and those where there wasn't. Constraints on archival research precluded the possibility of hand-picking a sample of settlements that might best illustrate the process. There was no guarantee of access or that useful materials even existed. The general dearth of information made this qualitative piece by far the most challenging part of the project. I spent most 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 institutions against communist encroachment. 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. The best reason to do this came from my publisher, who wryly informed me that books on Hungary were not best-sellers, and that without including other countries I had no hope of getting a paperback. Tempting as it was, I could see no theoretically 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 polity would not add variance that did not already exist within Hungary. Another frequent suggestion was that I include an in-depth narration 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 intellectually it had even less to recommend it than going cross-national. Such a narrative would not have revealed information that was not already available in more encapsulated form elsewhere.

The most potentially damaging criticisms suggested that I got Hungary all wrong. Some with country knowledge claimed that my findings merely reflected the fact that the region I had focused on had always been among the most conservative and Catholic in the country, and was thus not representative of Hungary as a whole. Quantitative people questioned some model specifications and my reliance on ecological data. I took these criticisms very seriously because I felt like if I didn't have Hungary nailed down I was doomed. Consequently, I devoted significant effort 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 confessional effects that I could not explore with materials that focused only on Catholic activity. On the quantitative side, I established that the ecological results were robust to many different specifications and corroborated any ecological inferences with comparable survey data. In the end I got the book contract, but not the paperback edition!

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 crunch numbers 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 researchers are not expected to have equal mastery over their cases, even those they investigate more thoroughly as part of a nested design.

3. If your research is primarily on one country, make sure you have sufficient within-country variation across subunits, over time periods, or across functional issue areas. Make really 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 asking 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.

I conclude by re-emphasizing the importance of starting with a good question. If your topic is truly compelling, you may be forgiven some minor sins, but no amount of methodological razzle-dazzle can compensate for a poorly posed problem.

Notes

[1] I am grateful to Andrew Bennett, David Collier, and Nick Ziegler for numerous helpful suggestions.

References

King, Gary, Robert O. and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Wittenberg, Jason. 2006. *Crucibles of Political Loyalty: Church Institutions and Electoral Continuity in Hungary*. New York: Cambridge University Press.

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 may be that multi-method research appears to offer the possibility of triangulating on a given research problem, 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 between 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 with a few well-chosen cases, or to marry "data set observations" to "causal process observations" drawn from focused qualitative research (Collier, Brady, and Seawright 2004).

The particular ways in which different methods should or can 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 (or two or four) instead of just one. A statistical analogy 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 analogy seems misleading, however, because applying different methods is not like drawing balls independently from an urn. In good multi-method work, various commentators suggest, the various 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 writers 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 Braumoeller 2006). Yet if where one starts affects where one ends up, the Pandora's Box of multi-method approaches is also 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, the selected ball and an additional ball of the same color are returned to the urn and the procedure is then repeated a large number of times. As Pierson (2000: 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 quite unpredictable, since we might end any 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 "single" method, the characteristic strengths of other kinds of research strategies can play 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 how different methods can inform each other, such that they 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 recent personal experience with conducting 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 Seawright 2007), but there has 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 inferences 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, in a different context, the "parallel 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 iterative 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 apologies for a lack of greater 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 orientation of this research before exploring several issues and challenges that arose in the course of conducting it. A near-consensus has emerged among scholars working in this area that oil and similar natural resources promote authoritarian-ism; at the least, a few recent analysts have suggested that natural resources may have more nuanced effects on the regime type. Yet some of the most resource-rich (if 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–among 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 in the relevant literature and by my familiarity with these several anomalous cases.

At the time I began this research, my disquiet about the claim that oil only hinders 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 displace non-resource taxation–then in such models one might expect more mixed effects of resource wealth, since resources could help ease the redistributive pressures democracy may sometimes impose on elites. The idea that resources could thus have mixed effects on the regime type matched intuitions that I had drawn from countries where I had done brief initial field visits, such as Botswana, Chile, and Venezuela.

In conjunction with reading the literature on the politics of rentier states and with further fieldwork, I began to develop a game-theoretic model to help me analyze these issues. There are always many analytic choices that go into the specification of an applied formal model. 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 underlying economic structure, what should be the relationship of resources to the non-resource sectors of the economy? These are just a few of the important questions that had to be answered before a model could be solved or its equilibria analyzed. The point I wish to make is that knowledge of 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 one, the literature on "rentier states" suggested that resources 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 lacked extensive "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 means of formalizing the contrast between rents and other sources of fiscal income in ways faithful to the claims of the rentier state literature (Dunning 2007, Chapter 3).

These brief examples may go to suggest that the process of developing a game-theoretic model can itself be a "mixed-method" process. Because analysts may draw on well-developed concepts or previous results in the field to stipulate core assumptions, developing a model may be considered a process that is both "inductive" and "deductive." This also implies, as mentioned above, that some of the distinctive strengths of "qualitative" methods, including especially tools for concept formation, can and often should inform the development of applied formal 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 inquiry, including case studies. If cases and concepts illuminated by previous studies help to motivate models, how can those models in turn be empirically "validated?"

A common and valid complaint about the merging of formal theory and case studies in many instances is that the case studies seem chosen merely to "illustrate" the theory. It might be useful to remember that this issue is far from limited to discussions of the interaction of qualitative and quantitative methods. Indeed, the point is reminiscent of what Skocpol and Somers (1980: 179) called in another context the "parallel demonstration of theory:" a form of empirical inquiry in which "the reason for juxtaposing cases is to persuade the reader that a given, explicitly delineated hypothesis or theory can repeatedly demonstrate its fruitfulness–its ability to convincingly order the evidence–when applied to a series of relevant historical trajectories."

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 if not sufficient 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 precisely the reasons 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. As Skocpol and Somers (1980: 191-2) also emphasize, the parallel demonstration of theory can avoid "repetitiveness" (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 theorist would call 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 vehicle for assessing the predictions of theoretical models empirically. Another point is that for those oriented towards formal work, case studies can not only provide evidence on the observable implications of a theory but can also help to motivate new models, an advantage of case studies that I found especially useful in my own work (Dunning 2007, Chapter 7).

An ongoing iteration between methods thus probably better characterizes most multi-method work than does the idea of one methodological "draw" followed by another. If one finding or methodological approach does condition the next, multi-method research hardly reproduces the non-ergodicity of a Polya urn process. Instead, 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 further on how apparently disconnected research strategies, such as concept analysis and game-theoretic modeling, may in fact complement each other in useful and possibly unexpected ways.

References

Acemoglu, Daron and James Robinson. 2006. *Economic Origins of Democracy and Dictatorship*. Cambridge: Cambridge University Press.

Bennett, Andrew and Bear Braumoeller. 2006. "Where the Model Frequently Meets the Road: Combining Statistical, Formal, and Case Study Methods." Manuscript, Departments of Government, Georgetown University and Harvard University.

Collier, David, Henry Brady, and Jason Seawright. 2004. "Sources of Leverage in Causal Inference: Toward an Alternative View of Methodology." In Henry E. Brady and David Collier, eds. *Rethinking Social Inquiry: Diverse Tools, Shared Standards.* (Lanham, MD: Rowman & Littlefield), 229-66.

Dunning, Thad. 2007. *Crude Democracy: Natural Resource Wealth and Political Regimes*. Book manuscript, Department of Political Science, Yale University.

Gerring, John and Jason Seawright. 2007. "Techniques for Choosing Cases." in *Case Study Research: Principles and Practices*. Cambridge: Cambridge University Press, 86-150.

Laitin, David. 2002. "Comparative Politics: The State of the Subdiscipline." In Helen Milner and Ira Katznelson, eds. *The State of the Discipline*. Washington, DC: American Political Science Association and W.W. Norton.

Lieberman, Evan. 2005. "Nested Analysis as a Mixed-Method Strategy for Comparative Research." *American Political Science Review* (August) 99: 435-52.

Pierson, Paul. 2000. "Path Dependence, Increasing Returns, and the Study of Politics." *American Political Science Review* (June) 94: 251-67.

Skocpol, Theda and Margaret Somers. 1980. "The Uses of Comparative History in Macrosocial Inquiry." *Comparative Studies in Society and History* 22 (2): 174-97.

Trends in Multi-Method Research: Sailing Ahead, Reckoning with Old Risks and New

Scott Siegel with Ariel Ahram, Julia Azari, Ashwini Chhatre, Bridget Coggins, Jana Grittersova, Matthew Ingram, Matthew Lieber, Claire Metelits, Tom Pepinksy, Andrew Pieper, Karthika Sasikumar, and Prerna Singh[1]

As the Institute for Qualitative Research Methods begins its seventh year, the experiences of past attendees illuminate some of the positive and negative aspects of engaging in multi-method research. Combining methods, when done well, can lead to substantial intellectual and professional rewards. Researchers make significant theoretical progress using multiple methods through the testing of alternative explanations that exist in our respective fields. The use of multiple methods also infuses our work with greater rigor and moves our understanding of the world forward. Yet, as the use of diverse methods is becoming increasingly popular, there are also good reasons to be wary about where their application is taking the researcher and the discipline in general. While we recognize the inherent limits in research that is purely method-driven, we also see how the inclination toward using multiple methods and their increasingly frequent application bring with them some intellectual trade-offs of which we should be more aware. In our zeal to apply different methods, we must also question whether we have enough knowledge and training to do one method, namely qualitative, well.

Our conclusions about multi-method research are drawn from our collective thoughts and experiences as we completed our doctoral dissertations. Approximately half of us have already completed our doctorates and are now or will soon be employed in political science or other disciplines at both liberal arts colleges and research universities. The rest have completed the data-gathering stage of our research and are currently writing up our results. We are a self-selected sample of students who attended IQRM in the past and wish to share our thoughts with the larger academic community. The views expressed below, therefore, may reflect some unintended bias, but we believe that they are widely shared by the wider academic community.

Given scarce time and intellectual resources, the decision to use multiple methods requires careful thinking and consideration of the benefits and trade-offs. The most frequent form of multi-method research was the combination of quantitative or statistical methodologies and qualitative methods. Aside from one of us employing a simple game, no one had chosen to include formal, game-theoretic methods in his or her dissertation. This interesting fact is most likely due to our fundamentally empiricist inclinations. Whereas both qualitative and statistical methodologies are comfortable with a reality that is often complex and full of multiple causal factors, formal modeling relies on a series of strong ontological assumptions and a degree of simplification that often does not fit the real world. On more pragmatic grounds, the use of formal methods usually requires additional training that is often not required by current graduate programs, while quantitative methods frequently are. Given scarce resources and time, investments in both formal methods and an additional one are prohibitively costly.

The most common practice of combining methods is selecting cases for qualitative investigation that are nested within a large-N quantitative analysis. Clear outliers or influential cases are identified and then qualitative methods are applied to trace the causal mechanisms involved. Some cases are also selected because they are important—ideal cases—to demonstrate the theory on the basis of a researcher's familiarity with the empirical sample under examination. Some of us go beyond descriptive statistics and employ various forms of regression and maximum likelihood analysis in order to determine which cases would be most fruitful for close historical analysis. Bridget Coggins, for example, selects cases for qualitative analysis to determine when and why some secessionist movements are recognized by the international community and others are not and remain subordinate to other states' justifications. Ariel Ahram also uses a nested design in his study of why the structure of an external threat shapes state development in the Middle East and Southeast Asia and these states' ability to achieve a monopoly of violence over their territory.

The application of quantitative methods is not limited to the use of regression or other typical quantitative tools. Some of us also use different forms of content analysis to test hypotheses that are usually amenable to only qualitative analysis. Here the line between qualitative and quantitative methods becomes increasingly blurred, especially for one of us who chose to use Ragin's fuzzy-set methodology. Through the use of content analysis and other techniques, Claire Metelits gained a better understanding about the treatment of the civilian population by rebel groups and her independent variables, such as the type of resources and the monopoly over their extraction, as well as the degree and type of variation that exists and how they are associated with each other. In order to explain if and why the American Left has become hostile to religious faith, Andrew Pieper employs both survey data and content analysis. In these situations, the qualitative analysis pushed or drove the use of quantitative approaches.

Combined methods are also used to compare and test different explanations in the literature for a particular phenomenon. For example, Scott Siegel combines different methods to determine which factors—domestic special interests, state strength, or national legal traditions—most likely explain why member states of the European Union violate supranational law. Both Matthew Lieber and Ashwini Chhatre employ multiple methods to assess not only which factors or conditions are most likely to explain political remittances by emigrants or successful forms of local collective action, respectively, but also to determine how these factors interact with each other. While none of our studies provide conclusive answers, the application of multiple methods moves the study of our chosen topics forward by helping us discard less useful theories and point us toward the more powerful explanatory ones.

One of the most fruitful results from combining different methods is maintaining a close dialogue between theory and evidence. For example, Julia Azari uses both methods to identify under what conditions US presidents claim electoral mandates. Karthika Sasikumar uses both qualitative and quantitative forms of content analysis to determine how the nuclear non-proliferation regime shapes the identity options of the Indian state. The combination of methods helped Prerna Singh see in new ways how regional nationalist identification and the closeness of electoral competition interact to affect social development outcomes. Similarly, Matthew Ingram uses a combination of quantitative and qualitative methods to see how political competitiveness and programmatic party commitments affected judicial politics in state courts in Brazil and Mexico. Jana Grittersova combines advanced statistical analysis and case studies to explore how financial interests and institutions influence decisions of governments with regard to exchange rate policies among Eastern European countries.

Finally, the combination of these two methods are sometimes used simply to increase the number of observations in order to generate a larger sample of cases and increase the confidence in our empirical findings, but in a unique way. Rather than simply relaxing the criteria that define our phenomenon of interest to increase the number of observations, close examination through qualitative analysis of the conditions that cause a particular outcome in a small number of countries leads one to inquire where else these factors exist and what is their impact in other settings. For example, Tom Pepinsky uses his extensive knowledge of a small set of Southeast Asian countries to consider what the implications of different political coalitions are for all authoritarian countries during financial crises, which were then analyzed through large-N statistical analysis. Tom and many others first use qualitative methods to examine the causal relationships between various variables and how they affect the outcome being explored and then investigate, through quantitative analysis, to what extent a select group of countries are a sample of a larger universe.

As illustrated by our chosen topics, the use of multiple methodologies did not significantly affect our decision to pursue interesting topics or normatively important questions. Nevertheless, because the deployment of multiple methods is especially conducive to the testing of alternative hypotheses that emerge out of the academic literature or for generalizing across significantly different political institutions, economies, and cultures, research questions that emerge out of an intimate knowledge of a country or region could fall by the wayside. Crucial issues relevant to policymakers or regional experts may not be considered as important or as vital as adjudicating various claims among rival theories. While em-ploying multiple methods is certainly a virtue on its own, it limits the scope of questions that can be effectively handled. In some situations, the application of one method is more appropriate if the research question is related to a particularly burning issue to which both the discipline and policymakers demand answers. Although the problem of avoiding policy-relevant issues is a concern for the discipline as a whole, using multi-method research to solve conundrums in the academic literature could become an exacerbating factor if the desire to use them precedes the formation of the research question.

Once multiple 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 improvement definitely remains. While many graduate programs offer established and cohesive sets of courses designed for statistical training, few departments offer any such training in qualitative methods. Committee members are not aware of which qualitative methods should be applied, the advantages of some methods over others, or of current developments in the field of qualitative methodology, especially regarding computer software. With one or two exceptions, the members of our dissertation committees either referred us to someone outside the department more familiar with these methods, mainly within economics, sociology, or communications departments, or left us on our own to explore the possibilities available and acquire the necessary skills to use them. As a result, the annual CQRM institute was the first and only time many of us were exposed to qualitative methods. A growing number of summer programs and other venues fills in gaps in graduate methods training to some extent as well, including the Political Methodology section of the American Political Science Association (PolMeth; PolMeth sponsors a summer meeting on advanced statistical methods; see http://polmeth.wustl.edu/); Empirical Implications of Theoretical Models (EITM; EITM sponsors a three-week summer institute on formal and statistical methods; see http://www.isr.umich.edu/cps/eitm/eitm2006/); the Interuniversity Consortium for Political and Social Research (ICPSR; ICPSR sponsors a variety of summer courses emphasizing statistical methods; see http://www.icpsr.umich.edu/); and American Political Science Association annual conference short courses and working groups sponsored by the Political Methodology section and the Qualitative Methods section. Their short, intensive nature, mainly during the summer, makes it easier to develop multi-method research designs and makes us more confident when applying them.

Many of us had very positive experiences when actually conducting qualitative research, especially those of us whose research involved travel and living in an unfamiliar location. For some of us, field research consisted mainly of interviewing a select number of government officials and digging through archives. Others delved deeper into their locales and did participant observation, generated surveys, or conducted open-ended interviews with a large number of respondents. These methods allowed us to become more familiar with the political, cultural, and historical environments of the places we studied. Irrespective of what we did in the field, however, we all share the view that we were relatively unprepared for what successful field research or the use of qualitative methods in general entails. The amount of time necessary to become familiar with our new surroundings, especially if they were non-Western, was underestimated, as was the time it takes to translate our findings from the field into usable forms of information. Although we all had quite ambitious projects and later learned how to scale them back, there was a sense of being lost at sea when doing our field research. We would have greatly benefited from more training with the actual mechanics of qualitative methods and the possible pitfalls to avoid.

Once completed, the reception of multi-method work appears mostly positive. Whether at academic conferences or when giving job talks, our colleagues generally welcomed the use of multiple methods, with the exception of a few political science departments that remain steadfastly committed to the legitimacy of a few methodological approaches. Even some fields remain dominated by uni-methodological research. For example, American Politics continues to be biased towards the combination of quantitative methods with formal modeling, while International Relations remains generally polarized between using this combination and the use of qualitative methods with a large-N statistical analysis. Comparative Politics seems to be the most open to multi-method research. Nevertheless, we can report that the methodological wars are slowly coming to an end, thankfully.

Political science departments conducting job searches appreciate candidates who can move easily from presenting rigorous and widely generalizable theories to discussions of particularly important or interesting cases. However, accomplishing this task is difficult, which is the source of some skepticism in the discipline about multi-method research. Given limited time and intellectual resources, mastering two or more methods is a challenge. Therefore, some more senior members of the discipline and hiring committees prefer job candidates who demonstrate a mastery of one method rather than doing slipshod work using two or more methods. In fact, many of us shared a thorough command of quantitative methods before embracing qualitative methods, which provided us with some increased credibility when on the job market. Once on the academic job track, the task of publishing continues to be shaped by a structural bias. Our belief is that there continues to be a shortage of journals that publish purely qualitative research or provide enough space for the inclusion of more than one method in a single article. In an atmosphere of "publish or perish," the perceived rapidity associated with publishing articles with the application of only one method, usually statistics, is made particularly alluring.

In summary, we believe that the use of multiple methodologies when conducting scholarly research is growing as the benefits become more familiar to young scholars. They are being used to create research designs that adjudicate among several alternative, established theories in our fields, allow us to make inferences that go beyond particular cases or observations from particular regions of the world, or develop more rigorous answers to important questions. Employing multiple methods in our research should never be one's first priority, however. In fact, the combination of different methods is not appropriate for all research questions, and we must remember that many topical and normative research questions exist that do not always require the use of more than one method. Second, because doing high-quality multi-method research requires so much investment in terms of time and intellectual resources, one clear trade-off is losing a close familiarity with the language, culture, or events that shape a particular period or place. In addition, there is already some concern that we underestimate the effort and time needed to carry out our multi-methodological research, while also being overconfident that we have a full command of a method's mechanics. Any skepticism or opposition to multi-method research projects reflects a concern about the quality of scholarship when different methods are combined and none are sufficiently mastered. As the number of multi-method research projects proliferates and increasing numbers of IQRM graduates begin to sail 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. Ahram, Graduate Fellow, Center for Democracy and Civil Society, Georgetown University, aia4@georgetown.edu. My dissertation, *Devolution From Above*, examines the phenomena of state-sponsored militias from the perspective of state formation. Nesting a case comparison of the Middle East and Southeast Asia within a statistical study of the entire developing world, I argue that the structure of external threat, itself path dependent on historical patterns of decolonization, inhibits or permits states to loosen their monopoly of violence within their territory.

Julia R. Azari, Marquette University (beginning in Fall 2007), Julia.azari@yale.edu. My dissertation, *Delivering the People's Message: Presidential Mandate Claims from 1929 to 2000* addresses the conditions under which presidents make claims to an electoral mandate in order to justify their actions. I approach the question using a combination of several research methods: simple quantitative analysis (i.e., charts, tables, simple significance tests); regression analysis; and qualitative case studies. By coding about 3000 presidential speeches, press conferences, and other communications for the inclusion of mandate claims, I have created a dataset that can be analyzed using quantitative methods. A subset of this dataset is used to demonstrate and explain the variation in the use of mandate claims across presidential terms beginning in 1929. The case studies have entailed primary source research at presidential libraries, with support from the Yale Center for the Study of American Politics and a Harry Middleton Fellowship in Presidential Studies from the LBJ Foundation.

Ashwini Chhatre, University of Illinois, Urbana-Champaign (beginning in August 2007), ashwini.chhatre@gmail.com. My dissertation, *Democracy on the Commons: Political Competition and Local Cooperation for Natural Resource Management in India* (Duke, 2007), explores the effects of democratic competition on local collective action in India. I use historical and ethnographic data for generating hypotheses as well as specifying context-specific measurements of the relevant variables for statistical analysis. I test the hypotheses on three data-sets covering indigenous and co-governance institutions managing forests and irrigation. The findings demonstrate that local communities are better at natural resource management than state agencies, but democratic competition constrains local collective action.

Bridget Coggins, Dartmouth College, Bridget.Coggins@Dartmouth.edu. My dissertation, *Secession, Recognition, and the International Politics of Statehood* (2006, Ohio State), studies the conditions under which secessionist movements are recognized as legitimate states by the international community or otherwise forced to remain subordinate to other states' jurisdiction. The first portion of the project is a large-N, survival analysis of great power recognition for secessionist movements (1931-2002). The second portion examines the causal mechanism behind recognition with two case studies, the dissolution of Yugoslavia and the wars of Soviet succession.

Jana Grittersova, Cornell University, jg288@cornell.edu. My dissertation, *Capture, Collusion and Consensus: Financial Interests and Exchange Rate Policies in Eastern Europe, 1990-2004*, examines the role of financial interests and institutional structures in the choice and sustainability of exchange rate regimes. I demonstrate that exchange rate regime choices depend on two dimensions: first, on the ownership structure and institutional variation of national financial systems shaped by the method of privatization; and second, on the different preferences of domestic incumbent and foreign financiers in exchange rate policy. I evaluate this argument using cross-sectional time-series econometric analysis (logistic regressions on panel data and duration models) of twenty-five EE countries between 1990 and 2004 and an in-depth examination of the cases of Bulgaria, the Czech Republic, and Poland.

Matthew C. Ingram, University of New Mexico, mingram@unm.edu. My dissertation examines state-level court performance in Brazil and Mexico from 1985 to 2005. The mixed-method research design combines (a) econometric analyses of time-series cross-sectional (TSCS) data across all states and (b) in-depth case studies of three states in each country, relying primarily on interviews and document analysis to trace the causal process. Building on growing literatures in comparative judicial politics and the effects of increased political competitiveness at the subnational level in nascent democracies, I study how competitiveness and programmatic party commitments shape the performance of state courts.

Matt Lieber, Brown University, matthew_lieber@brown.edu. My dissertation, *Remittances and Extraterritorial Politics: Overseas Voting Reform in Mexico and the Dominican Republic, 1994-2006*, uses the concept of political remittances to link remittance-sending communities to their home country institutions. Amidst a global trend to overseas voting, labor-exporting states have varying rules for extraterritorial participation. I conduct a structured-focused comparison of Mexico and the Dominican Republic, using process-tracing and actor interviews to pinpoint why each case adopted different rules. A large-N analysis of the global universe then tests case findings against established (unsystematic) explanations.

Claire M. Metelits, Washington State University (beginning in August 2007), c-metelits@northwestern.edu. My dissertation explains the change in behavior of rebel groups toward populations they claim to represent. Drawing upon fieldwork in Sudan, Iraq, Turkey, and Colombia, I explore shifts in the behavior of three rebel organizations: the Kurdish Workers' Party (PKK), the Sudan People's Liberation Army (SPLA), and the Fuerzas Armadas Revolucionarias de Colombia (FARC). Each group is examined in relation to three variables: the presence of rival groups, the need for resources, and the type of resources the group uses. I use process tracing and as well as crisp-set and fuzzy-set analysis to reveal that the presence of rivals is critical in determining changing rebel group behavior toward civilians.

Thomas Pepinsky, University of Colorado at Boulder (beginning in August 2007), thomas.pepinsky@yale.edu. My dissertation, *Coalitions and Crises: Authoritarianism, Adjustment, and Transitions in Emerging Markets*, studies how coalitions shape adjustment policy choice and regime survival during financial crises. In field research in Indonesia and Malaysia, I interviewed key decision makers and consulted newspaper archives to explain diverging adjustment strategies and regime trajectories during the 1997-1998 Asian Financial Crisis. I complemented these with four additional case studies from the Latin American debt crisis, and developed a quantitative test of my argument using all authoritarian regimes experiencing financial crises between 1975 and 1997.

Andrew L. Pieper, Kennesaw State University (beginning in July 2007), Andrew.Pieper@uconn.edu. My dissertation, *Competing Traditions? Religion and the American Left*, empirically tests claims that the American Left has become hostile toward religious faith, leading religious voters to abandon liberal ideology and the Democratic Party. I use three Left publications, *The Nation*, *In These Times*, and *Mother Jones*, to measure the American Left's attitudes toward religion from 1977-2000. Using content analysis, interpretive textual analysis, and NES data, I find that although the American Left has become more hostile toward religious individuals and groups, this enmity is primarily a response to, rather than a cause of, changing partisan preferences of some religious voters.

Karthika Sasikumar, Postdoctoral Fellow, Simons Centre for Non-Proliferation and Disarmament Research, University of British Columbia,Vancouver, BC, Canada, Karthika.sasikumar@ubc.ca. My dissertation, *Regimes at Work: The Non-proliferation Order and Indian Nuclear Policy* (Cornell, May 2006) argues that the nuclear nonproliferation regime, by constituting a range of possible identities for countries, facilitated India's forging of non-weaponized nuclear deterrence and its decision to go "formally nuclear." The regime's definition of the nuclear problem and its categorization of states into Nuclear Weapon States and Non-Nuclear Weapon States structured India's threat environment. The regime also served as a resource for domestic nuclear advocates. Secondary cases are the French and South African nuclear programs and the evolving counter-terrorism regime. Methods used are elite interviews, quantitative content analysis, and discourse analysis.

Scott Siegel, Naval Postgraduate School, snsiegel@nps.edu. My dissertation, *Law and Order in the EU: The Comparative Politics of Compliance* (Cornell, February 2007), explains why some member states of the European Union violate supranational law more than others. While levels of codification explain the cross-national distribution of violations, institutional obstacles, such as veto players, can prevent the quick settlement of these legal disputes. This thesis was tested using both quantitative and qualitative methods. First, I performed a statistical analysis of over 1200 violations of EU law and then process-traced a select number of violations in the United Kingdom and Germany. My findings suggest that only the combination of rational and sociological institutional approaches can explain the comparative politics of compliance in the EU.

Prerna Singh, Princeton University, prernas@princeton.edu. My dissertation, *Worlds Apart: A Comparative Analysis of Social Development in India*, combines case study research and statistical analysis in a "nested research design" in order to explicate the politics of pubic goods provision in developing countries, through a sub-national comparison in contemporary India. Based on survey research, an analysis of government documents, and 112 interviews with political elites, as well as participant observation of educational and health care facilities in a selected district in each of my four case study states, I argue that social development outcomes are determined by the interaction between the strength of subnationalist identification and the closeness of electoral competition.

Web Sites and Working Papers

Gary Goertz
University of Arizona
ggoertz@u.arizona.edu

One purpose of this newsletter is to provide information of use to section members. The newsletter already has sections devoted to giving abstracts of recently published articles and books of relevance to those interested in qualitative methods. With this issue, we would like to extend that to websites and working papers. Obviously, the internet is a huge virtual space. From time to time, the newsletter would like to alert section members to Web sites of particular interest to students of qualitative methods. I encourage readers to email me with suggestions of Web sites for future issues. In particular in this issue we would like to highlight two Web sites that post working papers (they both do much more than that, of course).

We all know that the road to publication can be a long one. Working papers are an important station along that road. They let people know about ongoing work often years before publication. They also give authors an opportunity to get feedback before submission or publication. The reader should be aware that these lists of working papers contain pieces that have subsequently appeared print and so these working papers are out of date.

COMPASSS

The COMPASSS Web site–http://www.compasss.org/Welcome.htm–has long had a section devoted to working papers. In particular, this Website is a place to get working papers that use Ragin's QCA or fuzzy sets methods or papers on these methodologies. Here we list working papers available on that site with 2005–2007 dates (the site has papers from earlier years as well).

2005

Caty Clement, "The Nuts and Bolts of State Collapse: Common Causes and Different Patterns? A QCA Analysis of Lebanon, Somalia and the former-Yugoslavia."

Peer C. Fiss, "A Set-Theoretic Approach to Organizational Configurations."

Gary Goertz and Jack S. Levy, "Causal Explanations, Necessary Conditions, and Case Studies: World War I and the End of the Cold War." NB: Forthcoming monograph edited by the authors, *Causal Explanations, Necessary Conditions, and Case Studies: World War I and the End of the Cold War*.

Frank M. Hage, "Constructivism, Fuzzy Sets and (Very) Small-N: Revisiting the Conditions for Communicative Action."

Carsten Q. Schneider and Claudius Wagemann, "Reducing Complexity in Qualitative Comparative Analysis (QCA): Remote and Proximate Factors and the Consolidation of Democracy."

Svend-Erik Skaaning, "Respect for Civil Liberties in Post-Com-