THE EMPIRICAL IMPLICATIONS OF THEORETICAL MODELS (EITM) WORKSHOP

REPORT

OF THE

POLITICAL SCIENCE PROGRAM, DIRECTORATE FOR SOCIAL, BEHAVIORAL AND ECONOMIC SCIENCES, NATIONAL SCIENCE FOUNDATION

together with

ADDITIONAL COMMENTARIES AND SUPPLEMENTARY DOCUMENTS

CONTENTS

ACKNOWLEDGMENTS

We thank Mark Jones and Joan Sieber for their suggestions and comments on earlier drafts of this report. We also thank former Political Science Program Director, Marianne Stewart for her assistance on many matters pertaining to the EITM Workshop. In addition to her comments and suggestions on this Report, Marianne's administrative practices provided a framework and afforded the necessary time so that the Political Science Program could expedite all matters pertaining to the EITM Workshop. The final thanks is to the participants of the EITM Workshop. Their commentaries and insights will have far-reaching consequences for not only political science but also for the other social sciences.

PART ONE: BACKGROUND AND EXECUTIVE SUMMARY

I. BACKGROUND

On July 9th and 10th, 2001, the Political Science Program of the National Science Foundation (NSF) convened a Workshop to seek ways to improve technical-analytical proficiency in Political Science by bridging the divide between formal and empirical analysis. The participants in the Workshop were senior scholars with research experience in various technical-analytical areas and proven track records in activities that have improved the technical-analytical expertise in various sciences. They have been editors and have served on editorial boards of leading journals. Participants were primarily from political science, but economics and mathematics were represented as well (see Appendix B).[1]

Formal analysis --- or formal modeling --- includes, among other things, deductive modeling in a theorem and proof presentation or computational modeling which requires the assistance of simulation. Empirical analysis --- or empirical modeling --- usually (but not always) involves data analysis using statistical tools. Both approaches provide significant scientific benefit to political science. At a most basic level, formal modeling assists in the "construction of valid arguments such that the fact or facts to be explained can be derived from the premises that constitute the explanation."[2] In contrast, empirical modeling shows the researcher where their model went wrong and leaves open the possibility that a more accurate model can be constructed.

A schism has developed between those who engage in formal modeling that is highly mathematical, and those who employ empirical modeling which emphasizes applied statistics. As a consequence, a good deal of research in political science is competent in one technical area, but lacking in another, that is, a formal approach with substandard (or no) empirical tests or an empirical approach without formal clarity. Such impaired competency contributes to a failure to identify the proximate causes explicated in a theory and, in turn, increases the difficulty of achieving a meaningful increase in scientific knowledge.

If one were to summarize in one word what bridging the divide between formal and empirical modeling means for the political and social sciences, that word would be

[1] The EITM Workshop was recorded and transcribed. The written transcript is available on the Political Science Program Web Site: http://www.nsf.gov/sbe/ses/polisci.

[2] R. Harrison Wagner, "Who's Afraid of Rational Choice Theory?" Typescript. (October, 2001), page 3. http://www.la.utexas.edu/~hw.

identification. The ability of a researcher to **identify** or parse out specific causal linkages among the many factors is fundamental to the scientific enterprise. Specifying a model that links both formal and empirical approaches alerts researchers to outcomes ***when specific conditions are in place*** --- and is also one of the best ways to determine an **identified** relationship.

The Empirical Implications of Theoretical Models (hereafter EITM) Workshop was conducted to suggest constructive approaches that the Political Science Program at the NSF could employ to foster identification in formal and empirical modeling. To these ends, EITM Workshop participants were asked to provide, prior to the workshop, a short commentary on the following issues:

1. Consider the factors contributing to the split between formal theory and empirical modeling. (This included the current status of the American, Comparative, International Relations, and Methods/Formal fields and subfields as well as other disciplines).

2. Discuss the need to bridge formal theory and empirical modeling and viable strategies for doing so in the discipline.

3. Discuss interdisciplinary avenues and extensions, which include academic and non-academic examples. These might include the work at academic institutions such as California Institute of Technology, and Carnegie-Mellon, and at non-academic institutions such as The Brookings Institution and The Santa Fe Institute.

4. Explore the role that NSF funding opportunities can play to advance the linkage of formal modeling and empirical modeling. What has proven effective in the past? Are there best practices in other disciplines?

5. Develop a coherent strategy for implementing the initiatives via a "Dear Colleague" letter from the NSF to the political science community. Modes of implementation might include:

    i) Infrastructure opportunities.
    ii) Annual meetings.
    iii) Graduate and/or undergraduate student opportunities.
    iv) Junior and senior faculty opportunities.
    v) Inter/multidisciplinary opportunities.
    vi) Other considerations.

II. EXECUTIVE SUMMARY

In both written or spoken commentaries, EITM Workshop participants recommended that the Political Science Program at the NSF address the technical-analytical divide between formal and empirical approaches in three priority areas:

- Education: Training and Retraining
- Dissemination of Knowledge: Conferences and Workshops
- Research: Establishment of Research Work Groups

Key suggestions concerning these priority areas were as follows:

<u>Education: Summer Training Institutes</u>

- Participants eligible to receive training and retraining should include graduate students, post-docs, untenured faculty, and tenured faculty.
- In the event there is more than one summer institute under operation, it is expected that linkages should be established between the various institutes to further the dissemination of knowledge to all participants and to the scholarly community at large.

<u>Knowledge Dissemination: Conferences and Workshops</u>

- Each individual workshop or seminar **must have a specific theme or problem** that allows for a variety of analyses which link formal and empirical approaches.
- Participants in these workshops and seminars may include a mix of graduate students, post-docs, untenured faculty, and tenured faculty.
- Where practicable workshop and seminar organizers are encouraged to establish linkages with the summer institutes and the possibility of organizing joint ventures.

<u>Research: Establishment of Research Work Groups</u>

- Each individual research work group **must have a specific theme or problem** that allows for a variety of analyses which link formal and empirical approaches.
- Participants in these workshops may include a mix of graduate students, post-docs, untenured faculty, and tenured faculty. That number shall not exceed 12 total members.

- Upon completion of the workshop, participants are eligible (as a team) to compete in the regular Political Science funding competition or future EITM research funding competitions.

More generally:

- The Political Science Program should fund up to $1,000,000 for these activities for fiscal year 2002.

- The Political Science Program must stipulate that all EITM proposals contain a **formal and empirical component**.

- The formal component and empirical component must be explicitly outlined. Formal components include (but are not limited to) game theory and dynamic stochastic modeling. Empirical components include (but are not limited to) applied statistical procedures and experiments. “Hybrid” techniques such as agent-based modeling are also welcome.

- The Political Science Program should encourage, when practicable, incorporating scholars and students from recognized and respected programs and institutions outside the United States in EITM activities.

- The Political Science Program should encourage, when practicable, interdisciplinary linkages.

PART TWO: PROBLEM DEFINITION, PROBLEM SOURCES, AND CURRENT ADVANTAGES TO REDUCE THE DIVIDE BETWEEN FORMAL AND EMPIRICAL ANALYSIS

I. PROBLEM DEFINITION

EITM opportunities for education (training), knowledge transmission, and research work teams are designed to bridge the gap between formal and empirical by addressing the factors that have produced that gap. In their deliberations, EITM Workshop participants were in general agreement that the separation was somewhat natural and is not confined to political science. The divide exists in other social sciences, including economics, where individuals specialize in either formal or empirical analysis due to their level of mathematical background and the type and years of training the substantive area or field requires. The divide also exists in the other sciences. It was noted, for example, that epidemiology is much more comfortable with empirical modeling. The primary epidemiology journal, The American Journal of Public Health (AJPH), does not usually publish articles that have substantial formal modeling. The major funding organization for epidemiological research, NIH, tends to support very few formal modeling projects.

Differences between formal and empirical approaches occur in intellectual outlook, skills, training, and research focus. In terms of outlook, formal modelers typically emphasize, in minute detail, linkages between concepts, whereas empirical modelers do not want to spend their research time parsing through minute details that may not add to their understanding. Formal modeling also requires analytical, logical, and mathematical modeling skills, while empirical modeling is inductive and, therefore, places emphasis on descriptive and statistical skills. Workshop participants noted that the intellectual investment needed for formal modeling is greater; it requires more mathematical knowledge than does empirical modeling to analyze a problem of interest. Training priorities differ as well. Empirical modelers devote their energies to data collection, measurement, and statistical matters, while formal modelers center on mathematical rigor.

These differences in outlook, skills, and training are reflected in distinct research practices and outcomes. For empirical modelers, model failures lead to emphasis on more statistical training or more sophisticated uses of statistics --- usually to "patch over" --- a model failure (see Appendix A). Formal modelers, on the other hand, deal with model controversies by considering alternative mathematical formulations but this is usually done piecemeal. The basic framework, such as expected utility, usually remains in place. The one similarity, however, between these two approaches is that both formal and empirical modelers tend to remain tied to their particular technique despite the warning signals evidenced in model breakdown.

II. PROBLEM SOURCES

The literature in political science consists of a proliferation of non-cumulative empirical studies usually without any formal component. Computing power has made it possible for more detailed, robust, and sophisticated data analysis than ever before, but this has become an end unto itself. The number of empirical modeling articles far exceeds that of articles that use formal models. More importantly, the number of articles that combine formal and empirical analysis is very small.[3] EITM Workshop participants singled out three leading sources for the current situation: compartmentalization, (under)graduate education, and career pressures.

A. Compartmentalization

Isolation --- ***compartmentalization*** --- of fields and sub-fields is the *status quo* in political science. Fields in political science, as reported by the American Political Science Association, include: American Government and Politics, Comparative Politics, International Politics, Methodology, Political Philosophy and Theory, Public Law and Courts, Public Policy, and Public Administration.[4] This current field and sub-field structure exacerbates the separation between formal and empirical models. For example, focusing on a question that is particular to American Politics increases specialization and, turn, discourages integrating approaches and theories that would best come about from studying a particular research question in many countries.

[3] This viewpoint is supported by an informal survey by EITM Workshop participant Carl Simon. He contrasts political science articles in the last twenty years. Starting with eighteen articles in two issues of the 1981 American Political Science Review (APSR) only one of these eighteen centered on formal modeling (6%). The other seventeen were purely empirical (94%) and none of the eighteen articles combined formal and empirical modeling. In the ensuing twenty years things change slightly in political science. In the sixteen articles in two issues of the 1998 APSR four articles were theoretical (25%). The other twelve were purely empirical (75%). There were no articles that combined formal and empirical modeling.

[4] For an extensive and important discussion of this issue and many others relevant to EITM, see Rebecca B. Morton, *Methods and Models: A Guide to the Empirical Analysis of Formal Models in Political Science*, New York: Cambridge University Press (1999).

Isolation of fields and sub-fields results, in conceptual redefinition and proliferation across fields.[5] This intensifies measurement and theoretical problems while producing regressive research practices, fiefdom mentalities, and outdated views of formal and empirical analysis.[6] One such outdated perspective about formal and empirical analysis is the assertion that these technical-analytical approaches are simply interesting intellectual enterprises that lack political and social relevance. This most basic form of misunderstanding about both formal and empirical analysis is only encouraged by compartmentalization. Why bother to model findings if one does not seek to generalize and predict in other areas?

In addition, the consequences of isolation between formal and empirical modeling can be found in problems of misspecification. Many formal modelers feel uncomfortable with powerful empirical concepts such as social norms, limited rationality, and psychological factors such as personality and identity.[7] The usual argument is that formal models are not meant to fit data, or should not be. While there is much to be learned from pure theory and abstract formal arguments, the formal modeling isolation reinforces distance from basic circumstances that these abstract models could help to illuminate.

Empirical modeling isolation, on the other hand, is equally guilty of not advancing scientific understanding when it fails to incorporate their "more complex and general assumptions" into a mathematically identified model with direct and testable implications. Instead "errors" or "confounding variables" that derail the inferential process are treated as statistical problems that require only statistical fixes.

In sum, EITM Workshop participants were in agreement that compartmentalization was not neutral in its effect. The effect is negative. It was proposed that one way to reduce the effects of compartmentalization was to separate political science into the study of domestic and international politics. Theory, data, and method would cover more general circumstances and lead to deeper understanding.[8] For the purposes of reducing the

---

[5] Consider the concepts of "Democracy" and "Power." For a discussion of the conceptual problems associated with Democracy see David Collier and Robert Adcock, "Democracy and Dichotomies: A Pragmatic Approach to Choices and Concepts," in Nelson Polsby (ed.), *Annual Review of Political Science*, Palo Alto: Annual Reviews (1999), pages 537-565. For a discussion of the concept of Power and a more general discussion of concept dimensionality see W. Phillips Shively, *The Craft of Political Research*, Upper Saddle River, New Jersey: Prentice Hall (1998, Fourth Edition), pages 27-36.

[6] Workshop participants pointed out that compartmentalization and segmentation has helped create a situation where a segment of the political science profession views formal and empirical modeling as one and the same.

[7] A good example of the consequences of formal modeling isolation can be found in psychology. Despite a growing literature in mathematical psychology, a perusal of the Journal of Mathematical Psychology reveals that mathematical modeling tends to be limited to the simplest of individual learning and perceptual phenomena.

[8] An abbreviated list of research questions that are not studied adequately because of compartmentalization are: political corruption, size of government, levels and types of taxation, economic growth and

formal and empirical modeling divide, the effect of reduced compartmentalization by substantive field would encourage integration between formal and empirical analysis.

B. (Under) Graduate Education

In an ideal world, political scientists should be educated to do research that incorporates five major components: 1) theory (informed by field work or some "puzzle"); 2) a mathematical model identifying causal linkages; 3) deductions and hypotheses; 4) measurement and research design; and 5) data collection and statistics. However, one or more of these components often is absent in political science research and, according to the EITM Workshop participants, the quality of formal and empirical modeling in political science is substandard.

There are at least two reasons for this state of research competency. One is that rigorous formal and empirical training is a somewhat recent development in political science. Another is that that are significant obstacles in the current political science training environment. The first obstacle is time. Students who desire training in both formal and empirical modeling will take longer to get a Ph.D. and most graduate programs do not have the resources to support students for more than four or five years. Consequently, students take the sequence of formal or empirical modeling classes but seldom both sequences. In addition to classes in formal or empirical modeling students must take classes in their substantive area. For students in comparative politics there are field work and language requirements. What normally is sacrificed, then, is either the formal or empirical modeling sequence. Taking a single course in formal and empirical modeling is not nearly enough to develop competency to do research.

The second obstacle to establishing formal and empirical modeling competency centers on the training itself. The economics discipline is illustrative. Economics graduate students are required to take one full year (usually) of mathematics for economists. This mathematical (and quantitative) approach is reinforced in substantive courses which typically are taught as an analytic science in a theorem-proof mode.

Mathematical (quantitative) competency in most economics graduate programs is demonstrated not only in these foundational courses, but also in qualifying examinations in the summer after the first year of coursework. Students must clear this hurdle before being allowed to proceed with their Ph.D. Political science also has qualifying examinations but they are usually at the end of all coursework. Moreover students are not required to take a qualifying exam in formal or empirical modeling unless that is considered one of their chosen fields of study. In fact, in some graduate political science programs students cannot make formal or empirical modeling a major field of study. The end result is that political science graduate students avoid developing basic competencies in formal and empirical modeling.

---

development, public debt, inflation, failed democracy, democratic stability, regime transitions, the rule of law, property and political rights, ethnic conflict, coups and revolutions, and terrorism.

C. Career Pressures

In many cases, younger or innovative scholars are not encouraged to master formal modeling, empirical modeling, or a synthesis between the two. The pressure for junior faculty to publish and earn tenure can discourage (re)-tooling and acquiring competency in both formal and empirical modeling. Whether having earned tenure encourages re-tooling is an open question. Tenure allows for risk taking and a longer-term view of one's research, it can also bring administrative demands, and a grooved research record that may be associated with risk aversion and complacency.

A discipline that provides few incentives for risk taking and re-tooling, but many for an assembly-line model of research production, is a discipline that imperils innovative theories and methodologies and, in turn, scientific breakthroughs. One could make the argument that EITM or initiatives like it are unnecessary because the unfettered marketplace of ideas expedites best scientific practices and progress. But, it is precisely because there are significant rigidities (training and otherwise) in the current academic setting (imperfect competition) which makes EITM-type initiatives not only necessary --- but imperative.

III. EXISTING ADVANTAGES

Despite the obstacles to bridging the divide between formal and empirical modeling, political science possesses several qualities which have the potential to reduce this gap. One important quality stems from a perceived weakness in formal political theory --- the lack of a general political equilibrium theory. William Riker characterized a general "political equilibrium" in the following way:

> [politics] involves the amalgamation of individual preferences into a social choice and subsequent enforcement of that result. At this general level, the goal of political theory is to identify the conditions for an equilibrium of preferences. Such an equilibrium is a social choice that the members of every sub-group in the society that are capable of bringing about a social decision prefer to any other alternative. This equilibrium is one the society will arrive at for certain, regardless of its particular institutions; and if by reason of some obstruction the society is deflected from it or forced to abandon it, the society nevertheless will return to it if the obstruction is removed.[9]

A big attraction of general equilibrium analysis was, and is, the formidable analytic power or traction it provides. However, since the late 1940s, research on the question of

[9] William Riker, "Political Theory and the Art of Heresthetics," in Ada Finifter (ed.), *Political Science: The State of the Discipline*, Washington, D.C.: The American Political Science Association (1983), pages 47-67.

a general political equilibrium has suggested that it does not, and will not, exist. It has been impossible to achieve due primarily to "the distribution of tastes in society."[10] As a result, politics and political science do not have a general equilibrium theory to facilitate standard solutions.[11]

However, many of the assumptions that accompany general equilibrium theory are questionable. In particular, "clean" solutions often require assumptions such as perfect foresight, hyper-rationality, common knowledge, standard discounting, and expected utility maximization. These concepts have been contradicted over and over again by empirical evidence. That is, political science cannot realistically assume some of the analytical simplifications that allow the researcher to derive solutions relating individual behavior to a set of political factors.

Consequently and somewhat ironically, this "problem," the lack of a general political equilibrium, means there are fewer impediments to adopting a new style of work emphasizing partial equilibrium. This approach would relax many of the standard general equilibrium assumptions, and rigor would be increased, not sacrificed so long as there was a commitment to merge formal and empirical analysis. Indeed, the thrust of recent work is that people behave in ways that are boundedly rational, and their motivations are better explained by work in cognitive psychology.[12] Since political science is not strongly allied to general equilibrium theory, there is (and would be) far less discipline-wide resistance to the very complexities that economists avoid. Political science, therefore, in bridging the technical-analytical divide, would also be able to "skip an intellectual generation" and link formal and empirical models with richer concepts such as framing, limited foresight, and learning.

Another quality is the potential for collaboration between those who do field work and/or study history and culture, and those who wish to combine formal and empirical work. These opportunities include analysis of political (social) science problems that deal with (among other things) multiple goals of citizens (with and without limited choices), the endogeneity of rules, and preference changes (including regime shifts).

---

[10] Ibid., page 51.

[11] For a discussion of general equilibrium theory, see E. Roy Weintraub, *General Equilibrium Analysis: Studies in Appraisal*, New York: Cambridge University Press (1985).

[12] Examples can be found in Arthur Lupia, Mathew D. McCubbins, and Samuel Popkin (eds.), *Elements of Reason: Cognition, Choice, and the Bounds of Rationality*, New York: Cambridge University Press (2000).

PART THREE: THE ROLE OF NSF: PROGRAM PRIORITIES

I. PRIORITIES

EITM Workshop participants recommended that the Political Science Program at the NSF address the technical divide particularly its nature and sources in compartmentalization, (under)graduate education, and career pressures, in three priority areas: education, knowledge dissemination, and research.

A. Education

To address the skills deficit in formal modeling, empirical modeling, and especially both, support can be provided for graduate training, post-doctoral opportunities, and mid-career re-tooling. Such support can include, but is not limited to, courses in formal and empirical modeling. For graduate students, funding could be provided for an additional year or two of graduate school to complete both formal and empirical modeling sequences. For faculty, support could be given to visit another department on campus or another institution.

Support can also consist of summer training institutes and training centers that are positioned to serve larger numbers of individuals while reaching graduate students and faculty who are in departments that cannot offer this training. These individuals become exposed to more experienced social scientists who combine formal and empirical analysis. The forms of exposure can vary, ranging from a summer (semester) to shorter-term lectures or workshops (one-week).

Support can be directed further to the development and application of short-course instruction templates. An EITM-type module might include the use of interactive teaching devices implemented through the use of various software packages. A number of modules in a variety of mathematical fields now exists and could be provided to faculty and graduate students. In addition, EITM Workshop participants argued that revamping undergraduate training in political science is perhaps the best way to solve the skills deficit.

B. Knowledge Dissemination

The transmission of advances in knowledge and research can be expedited by support for conferences and workshops, and educational and interactive web sites. EITM Workshop participants emphasized that such conferences and workshops should include:

- One problem or theme.
- Representatives from different theoretical (formal and field work) frameworks to explain the problem.
- Empirical modelers to design tests for the formal explanations.
- A large contingent of young scholars.

These conferences and workshops can encourage the comparison of the strengths and weaknesses of different conceptual frameworks, while also introducing people to new types of data and new ways to analyze the problem at hand. Web sites provide additional opportunities for the development and release of "course materials," and modules of conference proceedings, and new software programming techniques.

C. Research

EITM-related research activities can be supported in ways that provide linkages to the infrastructure needs of the social sciences over the next decade. This includes laboratories for experimental research. Shared facilities, for example --- funded in infrastructure competitions --- allow for formal and empirical studies, and also for pilot studies made with very fast turnaround to the scholars involved and the scholarly community at large.

The funding of laboratory experimental research promises a closer connection between the empirical and corresponding theoretical modeling effort. Unlike previous support, the scale --- number of subjects --- would be increased 10 times from 10 to 15 up to 100 or 150. This would facilitate richer experiments and also allow for data collection, which can be shared by many investigators, who could not afford such a facility at their own university.

Although the conventional practice of funding original research by individual scholars also continues, it is evident tht support must be provided to people who work as part of research team. In this way, scholars who are well-trained formal modelers can work closely with scholars who are well-trained empirical modelers on a specific research question.

II. INTEGRATION WITH OTHER DISCIPLINES

EITM-type collaborations in education, knowledge dissemination, and research can promote interdisciplinary interactions. For example, although work in economics has been instructive, the economic paradigm of full information and costless transactions is too narrow to be a satisfactory model for political and social behavior. EITM-type opportunities allow for recognition of such promises and problems and, in turn, for construction of explicit, richer models of the individual by developing cognitively realistic (or empirically verifiable) theories. Political scientists who use purposive, goal-seeking, or intentional behavior can make use of the small, but emerging, field of "behavioral economics." This field is empirical, while also employing the insights of psychology and, thereby, playing off the standard neoclassical assumptions. A new behavioral political science can, like behavioral economics, be more realistic and empirically based.

Other disciplines also could combine theoretical and modeling expertise with empirical and experimental expertise.[13] Research groups might include political scientists together with anthropologists, economists, sociologists, experimental psychologists, and computer scientists. Under the umbrella of EITM, truly interdisciplinary research work teams and interdisciplinary research networks can encourage new research orientations for senior members of the profession and expose younger members (graduate students and post-docs) to new ways of thinking that have not yet entered the standard curriculum.

[13] Examples of this type of interdisciplinary approach can be found in Steven Durlauf and H. Peyton Young (eds.), *Social Dynamics*, Washington, D.C. and Cambridge, Massachusetts: Brookings Institution Press and The MIT Press (2001) and, also, in Robert Huckfeldt and John Sprague, *Citizens, Politics, and Social Communication: Information and Influence in an Election Campaign*, New York: Cambridge University Press (1995). In addition, there are various institutions that have either a history of making such interdisciplinary research a central mission or have recently begun along this path. Such institutions include California Institute of Technology and Carnegie-Mellon. There are also non-academic institutions with a similar approach including Rand, The Brookings Institution, and The Santa Fe Institute.

PART FOUR: CONCLUSION

Significant scientific progress can be made by a synthesis of formal and empirical modeling. The advancement of this synthesis requires the highest possible levels of communication between the two groups. Formal modelers must subject their theories to closely related tests while, at the same time, empirical modelers must formalize their models before they conduct various statistical tests. The point is not to sacrifice logically coherent and mathematical models. Rather, it is to apply that same rigor to include new developments in bounded rationality, learning, and evolutionary modeling. These breakthroughs in theory will be accomplished with the assistance of empirical models in experimental and non-experimental settings.

How will progress be measured? There are several performance indicators, including the number of articles that use formal and empirical analysis in the major professional journals. Another measurable indicator is the number of NSF grant proposal submissions by faculty and graduate students (doctoral dissertations) that use both approaches. However, the one area that may be the most difficult to measure is improvement in the quality of knowledge. In this regard, the ramifications of merging formal and empirical analysis is a transformation of how researchers think about problems and whether they take intellectual risks in synthesizing the model and testing it. When they do, the primary achievement of EITM will be a better understanding of the political and social world, more accurate predictions, and ultimately the provision of solid information to policymakers whose choices can profoundly affect citizens quality of life.

APPENDIX

APPENDIX A: Opening Statement

APPENDIX B: Participant Commentaries

APPENDIX C: EITM Proposal

APPENDIX D: EITM Workshop Agenda

APPENDIX E: The EITM Dear Colleague Letter

APPENDIX A

Opening Statement for the EITM Workshop

By

Jim Granato
July 9, 2001

Thank you all for participating in this Workshop. To paraphrase Admiral James Stockdale: "who are you?" and "why are you here?" The answer to the first question is simple: you constitute the very best that political science and other disciplines have to offer. Your scholarship demonstrates a willingness to engage in work that is innovative and that meets the very highest of standards. In short, the way you analyze questions in your research makes you uniquely suited to address the issues that led to the creation of this Workshop.

The answer to the second question --- why are you here? --- is more complicated. You are here because there is a growing sense that political science has endured a technical separation, between formal and empirical analysis, for far too long. Indeed, this separation serves as a barrier to the scientific study of politics.

What is meant by the "scientific study of politics"? Among other things, the scientific study of politics requires building theoretically informed models that take account of confounding factors that may undermine inferences (betas), predictions (y-hat), or conducting policy simulations --- or some combination of all three.

Consider how the split between the two approaches undermines progress. First examine the risks associated with a strictly empirical --- read applied statistical --- approach. Assume that the empiricist's "theory" dictates that the empirical model contains more than one equation. If one were to use a rigorous standard, then the empirical model would need to be identified and, thereby, satisfy order and rank conditions. But, even if a model is over- or just-identified (and the zero order restrictions are credible), it is still possible that various parameter magnitudes constitute a result that undermine the entire theory (i.e., an indeterminacy in a model that says the opposite).

Unfortunately, empiricists would not know this, given their singular approach. Instead, they note the model is identified and would dutifully report the t- and F-statistics (log-likelihood), the size and sign of the parameters, and believe they have created something valid that advances our stock of knowledge.

Yet, this situation is not necessarily long lived since any ex-ante or conditional forecast using these "indeterminate" within sample parameter magnitudes would be inaccurate, even freakishly so. There is also the distinct possibility that the residuals created in this estimation are not iid. Of course, it is possible to "hide" the problem by applying some residual weighting technique, which can be done, and is done. More on that later.

On the other hand, had the empirical model been derived from a formal model in a fairly straightforward way, it would become clear that certain limiting conditions of various parameter values produced the inconsistency between theory and outcome.

Now, consider a strictly formal approach. Assume that the modeler devises an elegant model that, after much work, produces a single equation with a closed form solution. She also determines that an empirical test of the model (with actual data) is in order. The model is linear (in parameters and functional form) so the modeler chooses OLS. She runs the regression, and sure enough, this conscientious formal modeler finds the residuals are not white noise.[14] What is the modeler to do? Well, the specification took a good deal of effort (maybe months) to devise, so to keep the specification, the modeler weights the residual variance-covariance matrix and applies GLS. *And Voila!* The residuals are now iid, and like the empiricist above, the formal modeler reports the t- and F-statistics and shows that the theory (hypotheses) is (are) supported.

What's wrong with this picture? First, the non-random behavior of the residuals is a clear sign that the model (and theory) is (are) misspecified. No application of GLS --- no matter how powerful and efficient the technique used to weight the residuals --- can cure this. Such practice is simply incoherent since it makes no sense to "correct" the empirical model using information created by the misspecification in the first place. More importantly, why would it ever make sense to "correct" a model by relying on the mistakes the model created? The model is wrong. It is as simple as that.[15] In the end, this will be borne out again and again by out-of-sample forecast failures --- both ex-ante and conditional. Nothing is learned, and nothing is gained. There is no advancement.

As you can see, we have similar outcomes, starting with different approaches. Are these examples exaggerated? Are they caricatures? One need only to look at the discipline's most selective "A" journals to uncover the answer. The journals are replete with empirical patchworks --- such as the weighting of residual variance-covariance matrices --- that attest to the failure to portray accurately political phenomena. Both approaches, acting independently or carelessly borrowing from each other, are equally guilty. These practices are pernicious.

---

[14] In most applications of maximum likelihood procedures it is rare to see checks on residual behavior, the effects of outlying cases (i.e., logit), or collinearity.

[15] In one sense all models are wrong. The issue here is whether the error left over is random. Patching models using techniques such as GLS border on non-falsification since nonrandom residual behavior indicate a systematic flaw.

And this is why you are here. With your help, the Political Science Program seeks ways to take the lead in ensuring that current practices that are a consequence of this split become a thing of the past. There are many ways for this to happen, and indeed in some quarters it is already occurring. But, this is not just about technique. Rather technique is a vehicle, that appropriately applied can be used to reach our ultimate goal: a deeper understanding of political phenomena.

A word about the excellent commentaries is also in order. In their discussion about the issues at hand, workshop attendees noted that the split is natural. Others also noted similar patterns in other disciplines. It should be said that whatever the degree of pessimism or optimism expressed in the commentaries, certain themes do exist. First, there is a "problem" with current technical practice. Second, is the conviction that something can be done. Third, NSF can assist in this exercise.

So, how does the practicing political scientist --- the practicing social scientist --- who sees the utility of reducing the divide, or is at least interested enough to give it an honest attempt, alter the way they currently practice their trade? A better answer to this question is a central issue on the agenda before us for the next day and a half.

While it would be presumptuous to think this issue will be resolved in this workshop, progress can be made. Indeed, as many of you have noted, progress has been made in recent years. These relatively scarce works, showing a link between theory and empirics, are found in unpublished manuscripts, articles in various journals, and conference papers. For the most part, this research is motivated by a variety of subfield specific concerns. However, they also contain a link between theory and empirics suitable for much wider applicability. The hope of extending that accessibility --- the implementation of the Workshop recommendations --- has not been lost on those who are participating today.[16]

[16] Some of the same issues in the last two paragraphs are raised in the book "Rational Expectations and Econometric Practice," by Robert E. Lucas Jr. and Thomas J. Sargent, eds. University of Minnesota Press. 1981 (page xi).

APPENDIX B

Participant Commentaries (alphabetical listing)

1. Christopher Achen (University of Michigan)
(E-mail: achen@umich.edu)

2. John Aldrich (Duke University)
(E-mail: aldrich@acpub.duke.edu)

3. James Alt (Harvard University)
(E-mail: jalt@latte.harvard.edu)

4. Henry Brady (University of California, Berkeley)
(E-mail: hbrady@csm.berkeley.edu)

5. John Freeman (University of Minnesota)
(E-mail: freeman@polisci.umn.edu)

6. William Keech (Carnegie Mellon University)
(E-mail: keech@andrew.cmu.edu)

7. Richard McKelvey (California Institute of Technology)
(E-mail: rdm@hss.caltech.edu)

8. Rebecca Morton (University of Houston)
(E-mail: rmorton@uh.edu)

9. Carl Simon (University of Michigan)
(E-mail: cpsimon@umich.edu)

10. H. Peyton Young (Johns Hopkins University)
(E-mail: pyoung@brook.edu)

11. Dina Zinnes (University of Illinois, Urbana-Champaign)
(E-mail: d-zinnes@uiuc.edu

Comments by Chris Achen

June 30, 2001

Colleagues at the EITM conference:

In each of the prior generations of political scientists, overcoming stasis and planning the scientific future of the discipline has meant disseminating the newest research techniques. When *that* has been done, we have always said, *then* political science will be scientific. We have worked hard, and the dissemination has always been achieved. Indeed, each step made us smarter. But deep problems always remain.

Thus I prefer not to focus here on how political scientists might be trained to take formal theory more seriously and to acquire more advanced statistical knowledge. For important as those goals are, the truth is that in nearly all the good departments, these battles, while still consuming resources and still causing casualties, are in the final generation. The bitter complaints about formal theory "dominating" departments (with two or three appointments!) testify to the steady ebbing of money, prestige, and intellectual respect from non-quantitative approaches. The tide has turned.

Inevitably, the current generation of scientifically-oriented political scientists will achieve what they have fought for--if not in their lifetimes, then not long after. Respectable departments will be populated primarily by people who do formal theory and statistical methods, as economics departments are now. And the kind of important intellectual reorganizations and proper training of graduate students and undergraduates that Dina and Becky stressed in their remarks will become possible, even inevitable. Political science will be, if it is not already, a serious science.

Certain subfields of the discipline have already or nearly achieved this blessed state. They have great achievements to be proud of. Yet one can see in the current state of their knowledge the daunting problems that will confront us all in the near future. It is this deeper concern that I want to focus on—not the immediate technical issues that concern us, but the more profound challenge of grounding the new theory in evidence in a truly reliable way.

For the single most daunting fact about the coming victory is how much remains to be done even in the provinces it has conquered. That is not to say that the immediate agenda is unappealing or unworthy. Tying statistical models to formal theory is not easy, and it is important. At the same time, it's not hard to see how to do it in principle. The task will inevitably draw talent, and in fact, is already doing so (McKelvey, Signorino, Schultz and Lewis, Sartori, and others). It would be hard to overstate how important such work will be to the next round of political science advances.

At the same time, additional conceptual holes are opening up—methodological gaps that the new work exposes but that we have formulated only poorly, much less fixed. At the same time, one can see nascent solutions emerging--the founding of a field in the manner that Walras and Edgeworth founded modern economics. We need only the right incentives and resources to exploit what seems to me a once-in-a-generation opportunity.

Let me give some examples, beginning with game theory and its relation to data. In the following, I have borrowed good ideas and felicitous phrases from various colleagues; the dumb expressions are mine.

A good deal of game-theoretic work consists of toy models. A prime example is the theoretical modeling of international crises. We now have several such models, nearly all of which have the qualitative feature that they depend on incomplete information on the part of at least one side. If punishment can be meted out in subsequent crises, or if there are audience costs, or if the risk of accident rises exogenously over the time of the crisis, then we should see delayed agreements, bluffing, and wars no one expected, just as we do in actual crises. In short, these models have been deeply successful: At long last, we have begun to understand why crises occur. Though we don't yet know which of the postulated enforcement mechanisms is at work in most crises, we do at least know where to look. And more importantly, we have something to teach sophomores about incomplete information and how it causes crises and war.

Now suppose that we want to assess the competing models. We have good data on international crises from the BCOW project and a lot of experience analyzing them, since they have been widely used for other purposes. We could solve each crisis game for its equilibrium (to keep things simple, let us suppose that it is unique), and then connect to data by assuming that the utilities (or distributions of utilities) perceived by the players are functions of the data plus an error term. (There are a good many subtleties here: Do the players make trembling hand mistakes in an otherwise common knowledge situation known to the analyst and the players? Or are the players error-free, but the analyst lacks some knowledge known to all the players? Or do the players have private information unknown both to other players and to the analyst? Different statistical models are implied in each situation, and it is not hard to imagine many additional scenarios, some quite realistic.)

Some tying together of theory and data of this kind has been done already, and it is impressive, even eye-opening, as a theoretical exercise. Moreover, with experimental data, joint theory-data models of this kind can be quite helpful in doing actual statistical analysis.

My point here, however, is that, important as it will be to work out these statistical setups, applying them to real crisis data is a very delicate matter, as Ken Binmore remarked to me over lunch more than a decade ago. The crisis models are not meant to fit data, or should not be. They are pathetically distant from decisionmakers' circumstances in real crises, as very modest acquaintance with history will confirm. Instead, they are cute constructs meant to create an "ah-ha!" experience, which they do very well. One learns a great deal from this literature that case studies, psychology, and "realism" can never convey. But to uncritically fit the models to observations makes no more sense than digging through archives to seek out the historical evidence for one of Lincoln's homey stories about Illinois voters.

The same might be said about legislative modeling, where analysts are constantly discovering that predictions fail because deals have been cut in advance, or about the growing literature on Bayesian attitude change and voting, where actual people flop all over the survey instruments. The models are extremely helpful and insightful, and they ought to be fitted directly to data occasionally. But in my view, stopping there is a category mistake. Careful data analysis (too often not done!) can be relied on to turn up brutal specification errors when these models are confronted with real live human beings.

The fact, therefore, is that at present we have too little empirical work with which to discipline formal theory. The behavioral work too often ignores theory, and when it does not, its tests are rarely sharp or persuasive. Moreover, there seems little prospect for credible direct tests of the game theory models themselves. Experimental work is always useful, but is subject to the customary doubts about external validity, especially on topics like international crises. Thus even among the best-trained people we have, working in subfields that have drawn energy and talent, reliable knowledge is in short supply. Winning the disciplinary wars is fine; accomplishing something permanent is another matter.

How is it, then, that theory and data should relate in political science? Of course, this topic is far too daunting for any one lifetime, let alone a single memorandum. But let me to begin a conversation by saying that, in my view, the theory-data connection should typically be rather different than most quantitative political scientists now imagine. Because I am a methodologist rather than a theorist, let me focus on that side of the relationship in my remarks.

To begin, I believe that the relationship of social science theory to data should be looser than the common wisdom among theorists suggests. Social science is not much like the harder natural sciences, where theory often has to imply a well-verified and precise quantitative finding. By contrast, our best ideas are most often expressed in qualitative terms. Consider, for example, the proposition that, if a government enforces contracts but otherwise just leaves markets alone, then there will be a vast diversity of foodstuffs like bananas for sale virtually every day, even in rural Nebraska in the wintertime. That is an astonishing insight, of vast benefit to human beings, and it is clearly empirically true. But we do not verify it with regression equations.

The same might be said of the workings of the balance of power, the benefits of free organization of political oppositions, and the value to political stability of letting everyone vote. Each of these is a hard-won social science insight, still contested in some quarters, that has saved millions of lives and improved millions more. But we did not learn them with contemporary quantitative tools.

How did we learn them? Insight joined to political experience is perhaps the simplest answer. In consequence, there is a fuzziness and nagging unreliability about all of them. Contemporary tools ought to enable us to do better. But the point is that our tools need to be directed toward large qualitative generalizations of this kind, not toward imaginary analogies with the natural sciences. The constant coefficient estimates we so often strive for make no scientific sense for us.

What we need from the empirical side of this discipline, then, is thoroughly reliable *qualitative generalizations* with theoretical bite. "Democracies don't fight each other" may be one such generalization; "party identification predicts the vote very well" seems to be another. Both these propositions have set off substantial decision- and game-theoretical literatures. (Admittedly, both would be more helpful if we knew what "democracy" meant in the first sentence and "party identification" in the second, but progress is occurring on those scores, too.)

Neither of these two generalizations about political life came from prior theory. (Yes, Kant had proposed the second one, but almost nobody believed him, and his arguments had been forgotten until empirical researchers surprised everyone with strong evidence.) Both generalizations are excellent examples of important discoveries. And both demonstrate how empirical work often comes before smart theorizing rather than following it, a phenomenon familiar from the natural sciences. Kepler's laws preceded Newton and structured his theorizing; the surprising discovery that black box radiation arrived in discrete units led to quantum mechanics. In short, empirical research has a role to play that involves its own kind of imagination and creativity apart from theory. It is not simply the dim-witted, dwarfish varlet following the theorist around and washing up the glassware.

To assume that role more often, however, empiricists are going to have to think differently than they usually have. Gerald Kramer once wrote that doing theory is relatively easy; it's learning whether it is true that is hard. And he added, political scientists tend to believe the reverse.

Empirical work, the way too many political scientists do it, *is* relatively easy. Gather the data, run the regression/MLE with the usual list of control variables, report the significance tests, and announce that one's pet variable "passed." This dreary hypothesis-testing framework is sometimes even insisted upon by journal editors. Being purely mechanical, it saves a great deal of thinking and anxiety, and cannot help being popular. But obviously, it has to go. Our best empirical generalizations do not derive from that kind of work.

How to stop it? The key point is that no one can know whether regressions and MLEs actually fit the data when there are more than two or three independent variables. These high-dimensional explanatory spaces will wrap themselves around any dataset, but typically by distorting what is going on. They find the crudest correlations of course: education increases support for abortion, for example. In the behavioral tradition, that counts as a reliable finding. But no one knows why education is associated with that moral position (higher intellect discovering the truth? mindless adoption of elite tribal norms? correlation with something else entirely?), and that leaves open the possibility that abortion attitudes do not work the way the literature says they do. Getting rid of this cheap sense of "empirical findings" is probably the central task that empirical political research faces.

Indeed, among practicing Catholics, it turns out that the opposite of the conventional finding holds: More education leads to more respect for the pro-life view. This suggests strongly that education acts as a catalyst rather than a direct cause: It strengthens the coherence of one's attitudes with one's group ties and brings them into alignment, a process Henry Brady has written about. Either conversion or selective recruitment might be the main force at work.

From the NES time series, cross-tabs confined to the largest group, white Christians, shout the right answer at the researcher, while all the homoskedastic and heteroskedastic probits have mushed their way helplessly through confusing national samples with Christian Scientists, Jews, blacks, and Mormons, each with their own special circumstances and unique causal patterns. With a dozen control variables in a representative national sample, one cannot see much of anything. But with a few variables in the right subsample, one can see a pattern that gives a clue to *all* the groups. As one African-American colleague once said to me, "I hate dummy variables for race. Let me see what's going on in each group separately." Doing so may leave us still not knowing precisely how education works in changing attitudes, but it will prevent "explaining" what ain't so.

When we can, then, we may need to pick special samples where our chance to see the effect is at a maximum. Some of these special samples may have to be created. For example, few tools would add as much to our knowledge of the democratic citizen as a long-term rolling time series sample of voters. This might be done along the lines of the Panel Study of Income Dynamics in economics, perhaps as an extension to the National Election Studies or an alternate version of it. The survey questions would be designed to test both the Bayesian micro-models and the rational-expectations macro-models of how the electoral system works for both politicians and the citizenry. We could do much else, too, such as watching voters over time to see whether attitudes or group affiliations change when the two are in conflict over issues like abortion. Of course, it would be expensive, but you get what you pay for.

Meantime, the immediate task on the empirical side is to reduce the proliferation of non-cumulative studies and move ahead to better-grounded theory with the new and exciting tools we have. Computing power has made possible far more detailed, robust, and sophisticated data analysis than ever before. Increasingly, we do not have to rely on the implausible assumptions taught in econometrics texts or used in conventional maximum-likelihood and Bayesian estimation. But we will have to adopt a new style of work to take advantage of our changed circumstances and dramatically promising opportunties.

As an instance of the altered perspective I have in mind, I propose the following simple rule: Any statistical specification with more than three independent variables should be disregarded as meaningless. With more variables than that, no one can do the careful data analysis to be sure that the model specification is what s/he says it is. With just a few variables, powerful visual tools can really help us *see* and really help us *theorize*. The result of this rule would be more careful and appropriate choice of samples, and more attention to what the data really say, neither of which is characteristic of the current behavioral literature. Phony generalizations would be caught more often; truly reliable generalizations would have a fighting chance.

Some of these generalizations will have been suggested by theory: We will be searching under the lamplight. But others will come from the darkness unillumined by theory, and will give theorists a new conclusion to match up to their own qualitative findings. The relationship of theory to evidence is reciprocal, not monocausal.

Moreover, and here I may differ from some respected colleagues, I do not expect that this relationship will always be logically tight. In the most important verification steps, we will not literally be mapping formal models onto data. Instead, both the theoretical and the empirical generalizations will be primarily qualitative in character. Some of them will be virtually identical to each other and confirmatory in that sense. Others will differ but will have a mutually supporting character. Above all, I would hope that they would be more consequential than most of what passes for our theoretical and empirical generalizations now.

To be helpful in this enterprise, I believe, empiricists need a conversion to quite new ways of thinking. The implication of this new viewpoint for the methods subfield—what it teaches, what it publishes—are not at all clear at present. We need to think. NSF could play a critical role in giving established scholars a chance to create the new political science and in training beginning scholars in the new methodology that is coming along.

Indeed, profound rethinking will be needed if the new political science that melds formal theory with formal inference is to fulfill its current promise.

Undoubtedly NSF can help; this meeting is an excellent start. I look forward to mutual development of our ideas, and to discussing strategies for the future.

Chris

Christopher H. Achen
Professor, Political Science
University of Michigan

Comments by John Aldrich

The Nature of the Problem:

I suspect that most or all of us are agreed that statistical methodology in political science looks rather different from econometrics, in particular, and psychometrics, to nearly as substantial a degree. Surely all are agreed that the level of mathematical theorizing in this discipline is noticeably behind the levels of even applied, let alone pure theoretical economics. One reason for this state of affairs is simply that rigorous theory is rather newer to this discipline than to (especially) those two social/behavioral sciences. Secondly, thanks to the efforts of Achen, et al., it is even more recent that there has been a near explosion in the growth of rigorous methodologists in the discipline – including those who are making original contributions, rather than (as was common in Chris and my "generation") more often translating from Statistics or Econometrics grad-level courses than creating original methods for problems at hand in this discipline. Time will solve those two problems. The more difficult ones are:

1. Lack of agreement on the value of theory in the discipline, let alone broad agreement on what that particular theory is or should be (aka lack of a paradigm).

2. Lack of valuation in the discipline for getting the results done right (whether in deriving the appropriate theorems or in applying the appropriate methods), relative to having some new substantive insight. The payoff still comes, that is, from big ideas rather than carefully rendered results.

Last year at this time, I was on the APSA Strategic Planning Committee. By far the most contentious issue that we had to deal with was the "complete takeover of the *APSR* (and *AJPS* and *JoP*, as well as the APSA Council and officer corpse, etc.) by the rational choice – statistical methods conspirators." That is, to a noticeable segment of the discipline (including Mr./Ms. Peristroika) the concept behind EITM is massively mysterious – we are one, not two groups.

There was, in this set of dissidents but also among others, the belief that the leading journal of the discipline should present articles readable to all in the discipline. There was little recognition or support for the opposite perspective, that the leading journal of the discipline should be presenting the newest and most significant advances, which almost never are readable to anyone but experts in that small part of the research frontier.

We "compromised" by imagining a new journal, one that would publish review essays and the like. The publications committee of APSA picked up a similar line of thinking, with the Council then adopting the proposal. See http://www.apsanet.org/new/planning/picrecommendations.cfm.

In time, this MIGHT mean that the *APSR* becomes more like, say *Econometrica*. For here, that would have to mean a journal that recognizes the value of careful theoretical and empirical work and that therefore symbolically and practically offers value to those who do this sort of work. Indeed, from this perspective, the APSA is offering a high risk/return bet. If it pays off, the discipline's highest scholarly value goes to that sort of work. If not, we return to the present, where some of that sort of work is recognized but not given the highest accolades.

My personal position is that theory has always been empirically oriented and that orientation has improved dramatically in recent years. Its strength is not only where theory is most developed (which generally means in democratic institutional settings), but where data collections are the most dense (e.g., the debates over the role of committees in Congress and over the role of parties in Congress are theorized fairly precisely and adjudicated on the basis of reasonably exacting data analysis). Method, qua method, has advanced considerably, working on an agenda (Achen's in his chapter in the *State of the Discipline* to a large degree) that has been common to everyone, theoretically inclined or not. Now may well be the time to seek at least a subset of both theorists and methodologists who see it as their mission to write theoretically strong, methodologically sound papers, not unlike, say, Alisair Smith or Curt Signorino have been trying to do in IR.

On separation between theory and method:

Departments handle their graduate training and program structure quite differently with regard to theory and methods fields. When I was at Minnesota, for example, theory was one of the three subfields in the (Ph.D. exam-testing) field of philosophy. Here at Duke we just recently put it in as one of the three subfields in the (Ph.D. exam-testing) field of methods. Our proposal for requirements for eligibility to test in that area include (to quote from our proposal, with my inserts in brackets):

> 1) Stats I & II (PS222 & PS233) [a basic statistics course followed by the core econometrics course,] 2) Positive Political Theory (PS230S) and Game Theory (PS243S) [what I call the good cop – bad cop courses (i.e., user friendly introduction to positive, with proofs of only key things like Arrow, Black, followed by a reasonably serious introduction to game theory)], and 3) two selective courses both in either statistical analysis or formal analysis (that is, either two courses in econometrics, time series, maximum likelihood estimation, measurement theory, etc., or two courses in advanced game theory or social choice.) Petition to waive PS222 or PS230 needs to be approved by the field chair.

Conversely, it would be inappropriate to expect all who are good in theory to be good in statistics, and vice versa, for several reasons. First, there should be specialization and division of labor. Second, while likely true in general, it is certainly true now (and in the foreseeable future) that not all methodological problems are rooted in the same (or sometimes even any) theory and not all theoretical problems are appropriately studied via statistics. The goal, I believe, is that there should be the highest possible levels of communication between those who are the best and most original theorists and those who are the best and most original methodologists.

Factors contributing to the split between formal theory and empirical modeling:
Current Status of comparative government and politics

James Alt

7/1/01

I have to start with a disclaimer. My credentials as a contemporary comparativist are threadbare. For the last six years my focus has been on institutions, fiscal policy, and voting in the American states. Of course, comparing states is comparative politics of a sort, and indeed one encounters debates among students of American state politics that resemble those over which students of big-C comparative politics argue. In terms of pushing for particular approaches, I'm also not a big supporter of theoretical debates. I think ferment is good but you don't have to spend a lot of time talking about it. I've always preferred just to get on with empirical work and let readers decide among theories based on the (empirical) results they produce.

So we want to think about the split between formal theory and empirical modeling, and what's special about comparative politics. Thinking outward from what I do, my research often goes through five steps. It starts with a puzzling observation. I think much empirical research is case or problem-driven in this way. Given a puzzle, it's natural to ask "what's the theory?" What does verbal comparative theory tell us we know about this sort of puzzle? Is there a formal theory that leads to a specific prediction we can test? What does a test look like? What is the available data or evidence? So that's five stages to think about: a puzzle, verbal theory, formalization, testing, and data.

Now let's ask, what's special about comparative politics in terms of these five steps? There certainly was a time when the answer would have been "Jeez, the *data* is just so hard to get." I don't think that's really the problem any more, though, obviously, more data is generally better than less, and some sorts of reliable data are very hard to collect, especially in less developed countries.

I think it is true that in comparative politics the *puzzle* is often at a very macro level, dealing with collective rather than individual behavior. This is just the nature of the field. For instance, we want to compare revolutions, or the structure of state-wide institutions, and the units of observation are often countries, over epochs. This gives us big aggregation problems across individuals, space, and time. It also makes trouble for rational theories, since we have to ask whether the units have beliefs or preferences characteristic of individual actors. We could exhort comparativists to want to study different (smaller) things, but I don't think we'd get very far.

One thing that is very special about comparative politics is the ongoing (verbal) *theoretical debate* at its core. Most of what you need to know about this debate is in the (p)review of Lichbach and Zuckerman (eds.), *Rationality, Culture, and Structure*, available online at p. 8 of http://emma.sscnet.ucla.edu/apsacp/APSA-CP_Winter_1997.pdf. The three words in the title of this must-read book are answers to the questions "What should students of comparative politics study?" or "What kinds of things are our dependent variables?" Each has many proponents.

For our purposes, most of formal theory is rational choice theory. Rational choice models index individuals by their preferences and beliefs and solve out situations of strategic interaction among individuals for their equilibria on the basis of (mutual, constrained) optimizing. Consistency is very important. The unity of rational choice models lies in their relentless micro-focus on individuals' attempts to satisfy their preferences as much as possible.

But in all cases, rational is only rational conditional on beliefs, and the study of beliefs belongs properly to the study of culture as well as the science of cognition. Should we recommend a strategic approach to culture? Sure, there are times when strategic calculations determine whether or not we invest in learning something. Is this a sufficient way to study culture? Clearly not.

Those studying structure sometimes seem to say "bigger is better", but even when they don’t their macro focus is on things beyond the decision of a single strategic individual. We might ask for explicit assumptions about aggregation of individual preferences in deriving macro models, but that has not yet produced much of value. The rational and structural programs do meet each other: structuralists study social forces and movements but also institutions, and institutions are the most frequently-cited source of "constraint" in constrained optimizing that characterizes rational choice. Another source of constraint in rational models is the economy. Individual preference is often derived from economic interest, which is often measured with structural variables.

When we turn to *testing*, we find further internal divisions in comparative politics. A big debate that won't go away any time soon is the nature of the explanatory enterprise. For want of better words I call this debate "empiricism" versus "understanding". Put me among those who believe that when we explain something we offer a proposition whose truth or falsity depends (to some extent) on observational evidence. Others believe that you can only understand an action in terms of its meaning for the actor. This view of methodology is more common among proponents of "culture" for obvious reasons. Yes, I recognize that the words we use to describe what I call observational evidence indeed have shared socially-contingent meanings. However, I also believe that language is flexible and responds in some degree to a real world. To avoid a long digression into causation, laws, the meaning of words, and the psychology of perception (yes, I believe in framing effects) this debate I propose to leave alone.

Lurking behind this methodological question is the hoary old chestnut of "theory versus area studies". No one answer fits all here. Whether or not you do area studies is up to you. It depends on to what extent you define your interests by geography or region. Interactions between those who study countries' or regions' histories, cultures, or institutions closely and those who seek to generalize across boundaries have been incredibly productive -- when those interacting come to see the gains rather than the threat from exchange. Some like to talk about the payoff from cooperation in the study of Congress between modelers and "soakers and pokers". The patient evolution of a precisely-formulated and measured incumbency advantage in British politics out of the old "private vote" described by election observers is an example of the same happy synergy. Sometimes the issue in this debate is framed as "must hypotheses be theory-driven or can they just be interesting?" I myself think we're entitled to ask for both. "Do you need theory in comparative politics?" Yes, if you want to be able to use the word "anomaly".

Finally, *formal theory*. There are three things you often hear when comparativists talk about formal theory. One is to ask "did the model contain anything that could not have been said equally well in words?" While you might think this question is just carping, I think it arises frequently in comparative politics for a good reason. The focus on a macro puzzle and a formalization based on interactions among individuals often don't sit well together and it is hard to see how the empirical test actually bears on the model's predictions. If you prefer, so many extra assumptions and claims about what has been satisfactorily controlled or randomized are necessary to get to the test that the gain from the formalization gets lost in the shuffle.

A second thing you hear a lot is that the models are all about features of American politics and don't fit comparative (i.e., other countries) politics. I don't think this is really true any more. Compared to ten or twenty years ago, there is plenty of action in some areas of democratic politics more distant from American institutions: multiparty coalition formation and government duration being two notable examples. There is more work than ever before that has both elements of formal theory and empirical work in it: without working hard I can think of recent books by Cox, Huber, Tsebelis, Londregan, Laver/Shepsle (OK, some of these are strictly one non-US country rather than comparative) among others and significant articles by Diermeier and various co-authors, Myerson, and others. A major interdisciplinary initiative, heck more like an industry, is Persson and Tabellini's *Political Economics*. Throughout the book there is generally a formal economic model and a formal model of something political interacting behind every prediction of a range of variables spanning corruption, size of government, levels of several kinds of taxation, growth, debt, inflation, and more.

The third thing you hear is that formal theory does not yet deal with a wide range of concerns of comparative politics outside democratic institutions in developed societies. Probably there is quite a lot of truth in that. Again without working hard, topics like failed democracy and democratic stability, regime transitions and especially cycles of democracy and authoritarianism, politics in the absence of rule of law, unstable property and political rights, economic development and growth, ethnic political strife, and discontinuous political change like coups and revolutions strike me as among those where interest and indeed even some quantitative analysis has run ahead of satisfactory and useful formal theory. Remember, I didn't say there was no work. Yes, there's Przeworski, yes, there's Schleifer, and others. But one obvious thing (getting ahead of myself in terms of this memo) that the NSF could do is lean toward supporting promising theoretical work (even in advance of empirical work) in these areas.

What might make things better? The best I can do in general is to reiterate what was said by the six winners of the Nobel Award in Economics with whom Lin, Margaret, and I conversed (Alt, Levi, and Ostrom, 1999). The conversations produced a remarkably well-integrated and consistent theme: a chafing dissatisfaction with the microfoundations of much political and economic analysis. All six Nobelists suggested, in a variety of ways, that the standard neoclassical paradigm of full information and costless transactions was too narrow to be a satisfactory model for behavior, whatever its status as a normative standard. They converged on two directions for research to strengthen these foundations: (1) doing more, and better empirical work; and (2) building explicit, richer models of the individual by developing cognitively realistic (or empirically verifiable) theories.

In the former case, they called attention to the fact that experimentalists in both economics and political science had challenged other theorists as not explaining observed behavior, pointing out that the core of economic theory when applied to many political settings does not work as well as it does when applied to highly competitive institutions, including two-party competition. In the latter, they called for a set of more complex assumptions about human behavior, as enacted in diverse situations. Evolutionary processes, language, inherent violence, feeling, framing all stand outside the microfoundations of conventional rational analysis.

As far as the specific questions about the NSF's role and the talking points, let me mention three initiatives or directions that I think would be particularly helpful in comparative politics. First and foremost I would put improved training in methods, particularly experimental methods. Much of the micro-analysis that is at the heart of rational models is best tested experimentally. Improvements in technology and innovations in experimental setups have not just created the possibility of comparative survey experiments but of taking lab experiments into different contexts. The formation of the political methodology group, with its summer meetings and programs where younger scholars present research, revolutionized statistical practice in political science. Everyone goes, everyone learns, everyone presents, and the average ability level in the field rises. It seems odd to me that "political methodology" is restricted to statistical methods, though, and I would hope to see a similar organization evolve, with the same sort of NSF support the PMG had, over the next few years.

Second, I really would like to see more postdoctoral fellows. And I mean real postdocs, people straight out from dissertations working on projects under someone's supervision, just like in the lab sciences. A lot of the things we need to convey, like how to fashion a model out of a theory and how to fit statistical tests to your model, can be learned well (best?) by doing, first-hand. As the level of technical competence expected by a field rises, it seems reasonable to expect that a PhD will no longer contain everything the young scholar needs.

Last, I want to pick up on one of Becky's points, that in the long run we need to change undergraduate education. That's a great point, and implementing it would mean getting the NSF's DUE to understand our needs much better than I feel they do at present (wanna bring a DUE officer over for some of the discussion, Jim and Frank?). There are at least three dimensions to this. One is what Becky says, bringing out undergrads with a better sense of what political science does. Another is to be more innovative about recruiting potential scientists from undergrad studies in other fields into graduate work in political science. I don't know exactly what the NSF role would be, but we can talk about it. Finally, just doing better political science for undergraduates, giving them the instinct of taking questions to the data rather than looking for an article, would transform the way a more general public thinks about politics and perceives political science.

Factors Contributing to the Split
Between Formal Theory and Empirical Modeling

Henry E. Brady

Identification of the factors contributing to the split between formal theory and empirical modeling

Although there are good reasons to want to overcome the split between formal theory and empirical modeling, there are also good reasons for thinking that the split is so fundamental that it might be very hard to overcome. This split is a fundamental feature of scientific endeavor that goes back to the early days of the "Enlightenment" in late 16th and early 17th century Europe. It reflects the differences between Descartes' emphasis upon the need for mathematical demonstrations of theories (an approach most notably carried out by Newton and his successors such as Laplace, Maxwell, and Einstein) and Francis Bacon's emphasis upon the need for observation and experiment (most notably carried out by Robert Boyle and his successors such as Faraday and Kelvin). The differences between the rationalist Cartesian sensibility and the empirical Baconian approach run deep. Maxwell, we are told, added terms to his electromagnetic equations simply because symmetry demanded it – and the result was a fuller explanation of electromagnetism. Kelvin claimed that if you couldn't observe it and measure it with numbers, then it wasn't really science, and he improved our theories of thermodynamics by improving measurement. Both approaches, therefore, are very useful ways to advance science, but they are really different ways of thinking. An empiricist would be dumbfounded by Maxwell's temerity in adding terms to his equations simply to satisfy his craving for symmetry, and a rationalist would recoil at the need to observe and measure everything because it might lead to a rejection of "occult" forces like gravity or electromagnetism.

The split persists in physics and in many other areas of science. There have been some who have overcome this split such as Galileo or Enrico Fermi. But it has been said of Fermi that he was the last great physicist who was both an accomplished theorist and experimentalist. In fact, today's physics, certainly today's high-energy physics, is deeply split between mathematical theorists and empiricists. And the theorists persist in postulating occult forces that cannot be measured and the empiricists persist in studying phenomena (e.g., superconductivity at high temperatures) for which there are inadequate theories.

Another example where the split is wide is in population ecology where those who develop mathematical models ("mathematical ecologists") and those who collect facts and do statistical analysis ("natural historians") have been at war with one another. Each side has argued that the other contributes little to the understanding of ecology. Thus, there is something fundamental about this split that is related to the rationalist/empiricist split in philosophy, and it would be foolhardy to think that it can be completely overcome.

Turning to political science, we find the same split. If I may engage in a bit of caricature, I would describe each side as follows. Formal theorists tend to think that you have not explained anything until you have a mathematical model of it, and they venerate elegance, simplicity, and deduction from a minimum number of principles. They also like theories that cover a variety of different circumstances, favoring simple assumptions and theories with wide applicability. Theories are often considered proved, or at least supported, when unexpected predictions from the models turn out to be true in at least some instances. Truth for formal theorists emerges when theories seem to connect with the real world at a few points, although not necessarily everywhere, and the real force of their accomplishment is the unification of phenomena and the provision of elegant explanations. They strongly distrust empirical analysis on the grounds that it is merely correlational and that something more is needed – namely their theories. They also dislike the messiness of empirical research which often produces numerous disparate, hard to reconcile, and limited findings.

Empirical modelers tend to focus on explanations for concrete, specific situations, and they typically draw upon theoretical perspectives in an eclectic fashion as a way to justify (or suggest) the inclusion of numerous different variables that might explain the phenomenon under study. At their best, they worry a great deal about ruling out alternative explanations. Consequently, they do not want to be committed to any particular perspective because it might limit their search for confounding variables. Truth for empirical modelers emerges when alternatives have been ruled out and a number of factors have been identified that "cause" the phenomenon. They put great emphasis upon good research designs, careful conceptualization and measurement, and statistical methods. They distrust formal theory because it seems to limit the mechanisms that might explain the bloomin', buzzin' confusion of the real world and because it does not always produce predictions that can be easily tested.

Current Status of Methodology, Modeling and Statistics

There are still far too many statistical "modelers" who think that a regression equation or a likelihood function constitutes a model. And there are too many formal modelers who fail to test their model in any way. At the same time, my strong impression from the Political Methodology Group meetings and from panels at other professional meetings is that more and more people are trying to link formal modeling and statistical modeling. There is certainly a great deal of this in the study of Congress, and there are some notable efforts in the other fields as well, even in the study of mass political behavior – although formal models have been less fertile in that area. For example, in the study of legislatures, parties, referendums, political economy, public opinion, and voting, the spatial model has become ubiquitous, and it is often tested with real data in very interesting and innovative ways. Game theory models crop up in all sorts of situations such as the study of corruption in developing countries, deterrence in international relations, ethnic conflict, decisions to participate in protests, and so forth and some tests of these models are often performed. Certainly those methodologists who attend the Political Methodology Group meetings are repeatedly challenged to develop theoretical foundations for their empirical work. Sometimes these micro-foundations, however, are not rational choice models but event history models, social network models, psychological balance models, or difference equation models. Oddly enough, many formal modelers know little about these types of models, and one of the current problems in the discipline, I believe, is the equation of formal modeling with rational choice modeling.

Bridging Formal Theory and Empirical Modeling

Within our discipline, these approaches are often at odds because of fundamental differences in what is thought to constitute truth and progress in explanation, differences in basic ideas about what animates the world, and differences in training.

The differences in what constitute truth and progress in explanation have been alluded to above. Formal theorists feel good when they have a model that "puts all the pieces together" in an economical and elegant fashion. I have great sympathy for this feeling, and in one of my (small number of) attempts to do formal modeling, I found myself very pleased when I developed a dynamic spatial model of presidential primaries that really clarified (at least for me) the role of the media and the importance of strategic voting versus uncertainty reduction in producing "momentum" in primaries. But I remember a well-known empiricist remarking that the model seemed to add very little to his understanding of the phenomenon. That might have been a comment on my modeling skills, but I suspect it went deeper than that. Empirical researchers often feel that the amount of labor that they must invest to understand a model does little to improve their understanding of the phenomenon. I have often argued with them that a great deal can be learned about conceptualizing variables (e.g., voter uncertainty) and the possibilities that something (such as strategic voting) could explain something else (momentum in primaries), but they often reject these ideas. Part of the problem is undoubtedly a simple cost-benefit calculation on their part – these models are hard to understand given their training so that the benefits are often meager compared to the large costs of understanding the models. But a deeper problem is that some empiricists feel uncomfortable with the idea that one or two mechanisms described by a model could explain a phenomenon. They are not hedgehogs, in Isaiah Berlin's terminology, who know one big thing, they are foxes who know many things.

Formal theorists, on the other hand, like to provide explanations in terms of big things such as expected utility theory or median voter theorems. They get upset when their big things, like expected utility theory, seem to fall before empirical investigation. I remember hearing economists vigorously attack Amos Tversky at Econometric Society meetings in the mid-1980's after he had given a talk about prospect theory. At the same time, they seemed quite comfortable with Mark Machina's attempt to salvage expected utility theory by essentially assuming that it was a good local linear approximation to the truth. Tversky simply went too far for them. Now times are changing with the advent of behavioral economics, but certainly many formal modelers feel uncomfortable with this movement. Nevertheless, it might serve as a bridge between empiricists and modelers because it is bringing social norms, limited rationality, and psychological factors such as identity into formal modeling. There is, in fact, no reason why formal models should not include these factors, and their inclusion will certainly make some empirical modelers such as myself much happier.

The differences in basic ideas about what animates the world are also important. Formal modelers, whether fairly or unfairly, have been identified with either a narrow model of individual self-interest or a too general model of mere consistency in choice. And sometimes formal modelers make things worse by defending themselves in terms of the weak theory of rational choice which only assumes transitivity (or something even weaker) and then going forth and using the stronger theory which assumes narrow self-interest. Formal theorists have also been identified with a unitary actor model for nations, classes, or political parties that seems too constricting to some. And as noted above, formal modelers have been too quick to reject notions like social norms, social identity, and political ideology. Empiricists, of course, drive formal theorists crazy with their multiple, conflicting psychological and sociological theories and ever more complex notions of nations, classes or political parties. The result seems to be a welter of little findings that do not provide a very coherent picture. And there is the danger that as explanatory concepts multiply, social science explanations will sound like those of the Scholastics in which phenomena were explained in terms of the "nature" of the objects involved. Thus, we are told that the nature of Democratic party identifiers is to vote Democratic – not a very satisfying explanation of why some people vote Democratic.

Finally, there are differences in training. Empiricists often spend enormous amounts of time learning how to collect and weigh facts (through case studies, survey research, content analysis, event coding, etc.), and they worry a lot about concepts and measurement. As a result, they often have, perhaps undue, faith in the reality of the things they measure such as ideology, party identification, or social identity. And they often find that it is easier to explain something by inventing a new concept measured with a new scale from survey questions, than to use a more economical approach based upon existing concepts. Empirical modelers also learn statistical techniques that try to deal with the various ways that "error" or "confounding variables" can derail the inferential process, and they learn about research designs that can tame these problems. But they don't learn very much about the way that explanatory variables might fit together into coherent mechanisms. Indeed, some of their approaches to research design, such as randomized experiments, are meant to limit the operation of social mechanisms so that the impacts of one particular variable can be studied in isolation from all the others. The result can be the incredible profusion of single explanatory concepts that we find in psychological research that are not knit together in any coherent theory. Nothing seems to be tied together. Empiricists, of course, also try to develop models with many different variables in them, and they are not ignorant about problems of simultaneous causation, interactions, functional forms, expectations, and many other problems. But they typically try to find methodological solutions to these problems rather than building models of how they might arise in the phenomenon under study. The result often seems clunky and hard to believe.

Formal theorists, on the other hand, learn a great deal about mechanisms, and they fret about all sorts of nuances of rationality, belief, equilibrium, and stability. They believe that their models provide strong guidance about how social mechanisms work so that the prediction of particular effects provides good tests of their models. Empiricists, who are skeptical about the assumptions and often ignorant about the degree to which the models do rule out alternatives, get little solace from these results.

These differences might be lessened if we had a highly predictive set of models and/or empirical methods that allowed us to isolate phenomena. Consider thermodynamics, for example. In thermodynamics, theory and experiment worked together to further our understanding of gases. Before the statistical theory of gases was developed in the late nineteenth century, physicists knew that gases could be characterized in terms of the gas laws which related pressure, temperature, quantity of gas, and volume. The statistical theory of gases – a formal model – elaborated upon these laws by viewing a gas as a collection of molecules whose velocity is affected by the amount of energy in the system. Once this model was developed, it was easy to see that the molecules took up space (so that there would be a limit to how much they could be compressed) and that they would interact with one another because of molecular and atomic forces (so that they would tend to force one another apart). It was also easy to see that pressure and temperature were both the result of the actions of the energetic particles so that these two concepts could be reduced to a more basic one, namely energy. Based upon this model, predictions were made about departures that should be observed from the classic gas laws, and these predictions were confirmed by experimentalists through the careful isolation of systems and exact measurements of their properties. This success story has often been repeated in physics and other natural sciences, but there are not many of these stories to report in the social sciences.

Despite all of these caveats about the likelihood that we can make much progress, it would be useful if we made some attempts to bridge the gap. Clearly better training would be useful. Political scientists need to know more mathematics without a doubt. Formal theorists need to have better training in empirical work that goes beyond statistical training. They need to know much more about what it means to rule out alternative explanations and to explain something. Formal theorists know part of what it means to explain something (they know how to describe a model), but they typically know very little about testing a model. Conversely, empirical modelers know a lot about testing, but very little about how to formulate a model and they need to know more about this.

What Can NSF Do?

I think that a number of strategies would help matters. One is simply having this conference. I think that I have some ideas about the nature and cause of the gap between formal theory and empirical modeling gap, but I am looking forward to what others have to say. I think that we might be able to identify some concrete things that each side needs to know about the other. Of course, we might also find out that hedgehogs cannot become foxes and vice-versa. My hope is that we can produce some hedgefoxes and some foxhogs.

I have long believed that political science students are typically under-trained and that we need to improve undergraduate and graduate training. In addition, post-doctoral opportunities, in the right circumstances, would help a lot. Mid-career re-tooling might also help in some limited instances, but it is hard to change people at that point in their career unless the payoff appears to be substantial. Summer conferences like those run by the Political Methodology Group can be very important because they can set standards and keep people in the game. One of the problems with political methodology for many years was that its major figures got sidetracked into other pursuits – usually substantive research that yielded greater rewards in the discipline. The PMG has not completely stopped this (and I'm not sure it should have), but it has managed to keep many of us concerned and involved with statistical methods. We like to go to the meetings and show that we have not completely lost our edge, even if we are somewhat delusional about our ability to compete with younger scholars. Thus, a group that tried to bring together formal modelers and empirical modelers might help in this way by setting standards and keeping people in the game.

It is also possible that NSF could encourage graduate programs that produce hedgefoxes. Unfortunately, there are not many places that are strong enough in both formal modeling and empirical modeling to develop such programs, and there are even fewer places that have people who can bridge the divide between the two areas.

Final Thoughts

As I reread this missive, I worry that I am too pessimistic. As one who has argued for at least twenty years that we need more truck and trade between formal modelers and empirical modelers, I do not want to end on anything but a positive note. This is a very good time to be thinking about taking steps to bring both sides together. Both sides are now strong enough that they need not fear being swallowed by the other. Both sides have something to gain by working together – formal modelers need to produce empirical successes and empirical modelers need to show that they can make reliable inferences. Both need to respond to the "Perestroika" attack on formal and quantitative methods. Finally, both sides now seem to have enough tolerance for the other to insure that the result could be very fruitful.

Comments, EITM Workshop
John Freeman, U. Minnesota
Preface.
I. How deep is the divide? It is true that a decade ago, our functional forms were more likely to come from a chapter in an econometrics text than from rational choice theory. Many political methodologists knew little about recent developments in formal theory; most formal theorists knew little about advances in political methodology. Much of formal theory was motivated by “stylized facts” drawn from thick descriptions of politics; story telling was the method by which formal theory often was tested. Green and Shapiro (l994) and others were justified in raising this issue then. But, as the last part of Green and Shapiro’s own book shows(!), the situation changed, especially with regard to the study of micro political processes. In the last 10-15 years much work has been devoted to testing formal theories and to providing theoretical foundations for statistical models and estimators. The development and application of the random utility model is illustrative. Also, there now is a group of highly skilled young scholars who are achieving important syntheses of formal theory and methodology. For these reasons, the divide is not as deep as it once was.

II. The divide may be natural. There are well-developed branches of natural sciences that are largely experimental in nature. I am not sure whether a similar divide exists in these sciences, e.g., between theoretical and experimental physics, or whether physicists consider it problematic. Also, such a divide exists in economics. In fact, I believe that the Economics Division of the NSF has a separate panel for experimental work. There is conscious effort to bridge the divide in economics, of course (see below). The naïve view is that the experimentalists generate facts the theorists explain and also test the generalizations the theorists produce. This is a natural division of labor in mature (viz., increasingly specialized) disciplines. So, to the extent such a divide exists, perhaps it is a sign of the maturation of our profession.

III. What would a unified approach entail? As I understand it, methodologists’ functional forms would be derived from formal theory and their estimating equations would account for strategy, rational choice, etc. At the same time, formal theorists would incorporate probabilistic elements in their set-ups so that the error terms in the equations they estimate, distributional assumptions, etc. would flow directly from their mathematical models. In macroeconomics, for instance, this yields the so-called dynamic stochastic economy and set-ups such as the dynamic stochastic general equilibrium (DGSE) model. The rationales for introducing microfoundations are aesthetics, policy-relevance, and parameterization; incorporation of the stochastic elements into the formal theory requires mastery of dynamic programming methods (Ljungqvist and Sargent, 2000: esp pps. xxxvii-viii; Lucas, l987). A set-up along these lines in political science is Alesina et. al.’s model of the U.S. (l993). It is, in a sense, a dynamic stochastic political economy. Achen, Rivers, Mebane and others developed such models. My own efforts in this regard are in Freeman and Houser (1998) and Houser and Freeman (2000). Houser and I stress the value of such models for counterfactual analyses of such things as the welfare consequences of Presidents pursuing historically high levels of popular approval. In the spirit of the real business cycle theory, we argue that neither reduced form statistical models and nor pure formal models yield counterfactual insights that are as penetrating and useful as our synthetic political economic model.

IV. Caveats

a. There is a problem of observational equivalence (e.g., Sargent, l976). More than one formal theory may rationalize the same set of stylized facts and(or) fit the data equally well.

b. Macroeconomics seems to have gotten a lot of mileage out of representative agent models. My understanding is that its theories provide a justification for such set-ups. For instance, a perfectly competitive market behaves as if it was managed by a benign social planner. Hence one can employ the corresponding dynamic programming methods to solve for market behavior. I am not sure that political science has a comparable rationale for the use of representative agent models, at least for models of <u>democratic processes</u>.

c. If one uses nonlinear formulations of agent preferences familiar statistical methods are difficult to employ. This leads some macroeconomics to calibrate rather than estimate their models and to stress the ability of their models to mimic rather than fit observations of the economy. Needless to say, this alternative methodology is controversial (see, for instance, the symposium on calibration and computation in <u>The Journal of Economic Perspectives</u> 1996; see also Diebold, l998). Of course, there is a similar tradition in political science (Axelrod, l984). But, to my knowledge this calibration-computation tradition is not well-developed in either formal theory or political methodology. My point is that the pursuit of a synthesis might lead us into this debate because the models we construct cannot be estimated with conventional techniques. [The recent interest in Bayesian estimators (Jackman, 2000) conceivably could dovetail with computational approaches to solving formal theories (Jones et. al., l995)].

V. Formal theory's treatment of dynamics is very unsatisfying. I do not see how games with two or three periods (stages) help us explain time series findings about arms races, international conflict, governments' credibility vis-à-vis markets, and other important political processes. In fact, much of political methodology downplays or ignores dynamics. Political methodologists often assume events as independent across space **and time**. This might be justified in studies of vote choice and other micro processes. But it seems indefensible in studies of war, institutional design, and other meso and macro political problems.

Bottom Line. To me, the most important questions are: What scientific breakthroughs are likely to be made by the synthesis of formal theory and political methodology that would not be made otherwise? Are formal theorists better able to explain the stylized facts the political methodologists have given us than the methodologists? Do methodologists need formal theory to better calibrate their models, interpret their policy implications, and conduct counterfactual analyses?

Frank and Jim's Questions

1. Factors contributing to the split in comparative politics. Unlike other fields in our discipline, comparative politics (and IR) is(are) populated by people who are interested in conceptualization and thick description rather than theory and empiricism. The majority of comparative politics types at best have a very different (alien) notion of what constitutes scientific progress. Younger scholars therefore are not encouraged to master either formal theory or political methodology, let alone strive for a synthesis of the two. Rather, their advisors urge them to master of languages and cultures. Ideally, a scholar will master all these things. But, needless to say, only a handful of people in our profession have done this (Przeworki and Londregan are examples from two generations).

Exemplary (near) synthetic works in comparative politics, in my mind, include Laver and Shepsle <u>Making and Breaking Governments</u> (l996) and, more recently, Londregan's <u>Legislative Institutions and Ideology in Chile</u> (2000a; see also 2000b). Neither work achieves a synthesis of theory and methodology on the order of the DSGE. But Londregan's research on the design and workings of Chile's new democracy approaches this level of synthesis. Both works show how formal theory can explain empirical anomalies better than reduced form models, and that the introduction of formal theory into statistical set-ups can generate empirical results that might not otherwise be achieved.

NB. There remain in this field (as well as others) problems of data paucity, limits to experimentation, and the small n problem (Ragin, 1987) The first of these is being addressed by the new, NSF Comparative Elections Systems project. But the others are logical barriers to macro theory building at least as we conceive it here.

2.&4& 5. Viable strategies/NSF Role. We need courses—perhaps at Michigan in the summer—and at leading universities that teach the tools needed to achieve the synthesis. We also need books like Lucas and Stokey (1989) and Ljungqvist and Sargent (2000) that contain examples of synthetic political analyses. To my knowledge, few such courses and (or) books. (The new Morton book (1999) is a step in the right direction).

Whether a conference of people who have achieved a synthesis would succeed in generating sound proposals is unclear. As Frank might recall, in response to calls for research on democratization in Central Europe, we held a competition for conference proposals in the early 1990s. Three such meetings eventually were held at different sites. Faculty and graduate students attended. I do not know if many valuable projects emerged from these efforts. Several members of our committee told me this experiment did not produce much in the way of new and important research. Because of this experience I favor post-doctoral fellowships for selected faculty and graduate students rather than conferences.

3. Interdisciplinary efforts. I am curious to learn if our colleagues believe their institutional setting facilitated recent scientific breakthroughs, e.g., the idea of quantal

response equilibrium (McKelvey and Palfrey, l995) that lies the heart of Signorino's (l999) synthesis of formal theory and political methodology.

6. How to measure progress. It is important to emphasize criteria other than aesthetics. I know that some people feel the repeated calls for microfoundations at our summer methods meetings are driven by this concern rather than any vision of what a synthesis might yield in the way of scientific advances. The challenges are to convince (1) formal theorists that political methodologists have produced facts they must explain and also meaningful tests of their theories and (2), political methodologists that by incorporating formal theory they can solve important estimation problems and extract deeper and more useful insights from their results in ways and to an extent that are not possible by relying solely on existing estimation techniques and reduced forms.

7. Scholars who have and will met these challenges. Among the young people who come to my mind are Nolan McCarty (with Rothenberg, 2000), Curt Signorino (l999), and John Londregan (2000a,b).

References

Alesina, A., H. Rosenthal, and J. Londregan (l993) “A Model of the Political Economy of the United States” American Political Science Review 87: 12-33.

Axelrod, R. (l984) The Evolution of Cooperation NY: Basic Books.

Diebold, F. (l998) “The Past, the Present and the Future of Macroeconomic Forecasting” Journal of Economic Perspectives 12(2): 175-192.

Freeman, J. and D. Houser (l998) “A Computable Equilibrium Model for the Study of Political Economy.” American Journal of Political Science 42(2): 628-660.

Green, D. and I. Shapiro (l994) Pathologies of Rational Choice New Haven, CT: Yale University Press.

Houser, D. and J. Freeman (2000) “The Economic Consequences of Approval Management in Comparative Perspective.” Unpublished manuscript.

Jackman, S. (2000) “Estimation and Inference via Bayesian Simulation: An Introduction to MCMC” American Journal of Political Science 44(2): 375-404.

Jones, B., B. Radcliff, C. Taber, and R. Timpone (1995) “Condorett Winners and the Paradox of Voting” American Political Science Review 89: 137-144.

Laver, M. and K. Shepsle (l996) Making and Breaking Governments NY: Cambridge University Press.

Ljungqvist, L. and T. Sargent (2000) Recursive Macroeconomic Theory Cambridge,MA. MIT Press.

Londregan, J. (2000a) Legislative Institutions and Ideology in Chile NY: Cambridge University Press.

_____ (2000b) “Estimating Legislator’s Preferred Points” Political Analysis 8(1):35-56.

Lucas, R. (l987) Models of Business Cycles Yrjeo Jahnsson Lectures. NY: Basil Blackwell.

____ and N. Stokey with E. Prescott (l989) Recursive Methods in Economic Dynamics Cambridge, MA. Harvard University Press.

McCarty, N. and L. Rothenberg (2000) “The Time to Give: PAC Motivations and Electoral Timing” Political Analysis 8(3): 239-260.

McKelvey, R. and T. Palfrey (l995) “Quantal Response Equilibria for Normal Form Games” Games and Economic Behavior 10: 6-38.

Morton, R. (1999) Methods and Models: A Guide to the Empircal Analysis of Fromal Models in Political Science NY: Cambridge University Press.

Ragin, C. (l987) The Comparative Method Berkeley, CA: University of California Press.

Sargent, T. (l976) “On the Observational Equivalence of Natural and Unnatural Theories of Macroeconomics” Journal of Political Economy 84: 631-640.

Signorino, C. (1999) “Strategic Interaction and the Statistical Analysis of Political Conflict.” American Political Science Review 93(2): 279-298.

Symposium on Computational Methods [In Macroeconomics] (l996) Journal of Economic Perspectives 10(1): 69-120.

Empirical Implications of Theoretical Models Workshop
July 2001

Bill Keech
Carnegie Mellon University

Formal theory and empirical work

First, I do not think that a "split" between modelers and empiricists is new, and to some extent it can be part of a healthy division of labor. Quantitative research in political science was initially empirical, as reflected in the "behavioral revolution." Empirical research methods in the form of statistics became part of most graduate programs in the 1960s, but any idea of theory in the sense of deductive models was pretty rare (though not unknown) in the discipline in that decade. The idea that empirical statements could and should be tested systematically (even with cross-tabs) was important progress. (Political science spent plenty of decades without realizing or implementing this fundamental idea.) Scientific theory in political science was initially very inductive.

Deductive theory became more prominent as more political scientists were exposed to or trained in economics. Interest in economics was stimulated by Anthony Downs, and by Mancur Olson, whose modeling was very accessible. Olson's conclusions about collective action were at odds with the dominant pluralistic idea from David Truman that a shared attitude was sufficient to produce an interest group. Even more than Downs, Olson gave us something to think about - and a demonstration that deductive theory could explain stylized facts about the real world.

A major reason for the continuation of the "split" in my view is that the intellectual investment needed for formal theory at a level to get a payoff is greater than that needed for quantitative empirical work. It just takes more math to get off the ground in theory than it does in empirical work.

I think that the influence of economics has been basically constructive, but double-edged. On the positive side, here was a mature discipline focused on the study of the aggregation of preferences. Simply importing economic concepts and models into the study of politics could make considerable progress. These concepts and models were also part of General Equilibrium Theory, which I think is the greatest achievement of the social sciences.

On the negative side, economics was based on simple assumptions about human motivation that were more problematic for the study of politics than for economics. Economics was not strong on the question of where preferences come from, nor was it easy for economists to recognize that there are limitations on the capacity and inclination of individuals to gather and process information. Political scientists who incorporated the possibility of goal-seeking or intentional behavior took or were given the label of "rational choice" modelers, which continues to be unfortunately divisive.

A division of labor between modelers and empiricists is not inappropriate. Either one takes a significant investment. I think literacy in both is important for anyone getting a Ph.D. in political science. Sure, it is desirable for people to be strong in both, and to do both in their scholarship. But this is demanding more than might be expected in an era of specialization. Modelers should be literate in empirical methods and in empirical work. Empiricists should be literate in theory. The two should speak to each other informally if not formally. Green and Shapiro do something constructive in confronting theory with empirical work, but the tone of their book is unfortunate because it is so hostile to theory.

A role for normative issues in theoretical and empirical political science

Economics is a normative as well as a positive and an empirical discipline. Political science is positive and empirical too, but it has a different normative tradition than economics. The normative heritage of political science should not be forgotten or ignored. An example of an article that makes good use of the normative heritage is Canes-Wrone, Herron and Schotts in the July 2001 AJPS. This article has a normative topic: executive leadership and "pandering." It is theoretical in that there is a formal model with information, beliefs and signals. It is empirical only in the sense that there are three nuanced case studies to illustrate the relevance and applicability of the concepts. True leadership, pandering, and fake leadership are operationalized in the context of informational game theory. For something like this topic, I think that rich case studies are probably all we can hope for to give empirical referents and meaning.

American politics

Two of the best books on American politics I have read recently are by Barry Ames on Brazil and by John Londregan on Chile. Yes this is American in the sense that it is about politics in the Americas, but it is also like American politics in that both books are about preferences and institutions in the study of national level domestic politics. American politics has been the most arrogant of the area studies because it has been more theoretical and quantitatively empirical than the study of other countries usually is. The more books like Londregan's there are, the less we will be able to focus on the study of the US as the leading setting for the advanced study of politics, and the more we will know about domestic political processes everywhere.

Strategies for generating puzzles

Let me suggest two generic strategies for generating puzzles, the solution of which may lead to progress. One strategy is to maximize the exposure of apparently disjoint but mutually relevant streams of scholarship (like theoretical and empirical work) to each other. Keep them talking to each other. Make them confront each other's scholarship. This is something that NSF can do with conferences that it sponsors. The hope is that each side can learn from the other and will adjust their work so as to do so.

The other (not unrelated) strategy is to encourage the search for micro-foundations of aggregate phenomena. I am still impressed with the progress that was made in macroeconomics by the effort to make theories about aggregate behavior compatible with what was known (or thought) about individual behavior. Macroeconomics did this with what political scientists would think is a narrow and unrealistic conception of individual behavior. Another strategy is computational modeling of simpler kinds of behavior that can lead to emergent system-level properties, as in Bendor and Ting's model of turnout.

A new behavioral political science

Economics went far with unrealistic assumptions about rationality, information processing and foresight. There is now a growing field of "behavioral economics" that is empirical and uses insights of psychology, and plays off of the neoclassical assumptions. Behavioral political science came before formal modeling, while behavioral economics came after. A new behavioral political science can, like behavioral economics, be more realistic and empirically based. Unlike the first behavioralism in political science, it would play off of formal theory.

Roles for the National Science Foundation

What should NSF do? NSF should encourage literacy and awareness across related disciplines, and across theory and methods within political science if not other disciplines. It could do this with conferences of carefully chosen people.

I have two proposals for an NSF role in training, one predoctoral and one postdoctoral. First, NSF could have some fellowships that are awarded to individual advanced graduate students to help support the deeper and broader training that some of the other memos called for. I suggest that these fellowships be for one or two years of tuition and stipend support within their own graduate program, if it is approved by NSF for such training. There are dual incentives involved. The incentive of getting an NSF advanced training fellowship would be an important prod for graduate students to broaden and deepen their theoretical and empirical training. The incentive of having a graduate program that is approved for advanced theoretical and empirical training would prod departments to develop and maintain programs that would be beneficial for all of their students and for graduate training in general.

My postdoctoral proposal is for young scholars who have already achieved tenure, and demonstrated their promise with concrete achievements in either theoretical or empirical work. These post-tenure awards would be usable at NSF-approved programs as in the first proposal above, and be designed to strengthen either the theoretical or empirical training of scholars who had demonstrated distinction and promise in the other of these areas. Like the predoctoral fellowships suggested above, this would provide dual incentives for individuals seeking the awards, and for departments seeking to be eligible to receive the awardees.

Comments by Richard McKelvey

The focus of this workshop is the so called empirical/theoretical split in the political science profession. So I start with my understanding of the nature of this split, and then consider whether it is something that the NSF can or should be trying to rectify.

As others have already noted, there is certainly a split in terms of division of labor. With some notable exceptions, most researchers can be categorized as either theorists or empiricists, and tend to do either all theoretical work or all empirical work. This split between theorists and empiricists is a natural one. It exists in other disciplines as well, and is fostered to a large degree by the different skills that are necessary for each task. The theoretical enterprise is primarily deductive, and the hence attracts researchers whose skills are primarily analytical, logical, mathematical modeling skills. The empirical enterprise is primarily inductive, and attracts those with skills that are descriptive, statistical, and who are willing to focus in great depth on particular applied problems.

In terms of the body of research in political science, there is a lot of empirical research that seems purely descriptive and unmotivated by theory, but there is also a substantial body of empirical literature that is motivated by theoretical questions. Focusing on the later literature, my impression is there is not so much a split as that theorists are frequently not very convinced that empirical studies have adequately tested the theoretical models that they are based on. There are couple of reasons for this. First, the world that theorists look at is necessarily abstracted and simplified, focusing on the effects of certain variables at the exclusion of others. In the real world, it is seldom that one can find empirical situations where only the variables of theoretical interest are active. So econometric techniques, sometimes of questionable validity, must be used to introduce controls. Also, there are frequently problems measuring the variables of theoretical interest (utilities, beliefs) in natural settings. Thus it is frequently hard to get the data that is needed to test theoretical models. So empirical work encounters difficulties in the implementation, because of problems with operationalization of variables, and the inevitable econometric questions concerning uncontrolled confounding variables, whether the independent variables are truly exogenous, etc. This is partly because the theories are too primitive to be able to address the real world situations that empirical researchers are interested in and partly because empirical studies do not provide natural experiments.

The above difficulties are inherent in the empirical enterprise, and are more severe in the social sciences than in the natural sciences, because in the social sciences it is usually impossible to run controlled experiments in the real world. It is precisely this need to design experiments that control for variables that are not of theoretical interest that has led many theorists to embrace laboratory experiments as a more suitable avenue to test theoretical models than empirical studies. I notice that there has been no place in the program devoted to laboratory experiments in political science, and little mention of it in the other participants comments. But I feel that laboratory experiments provide one of the more promising directions to test theoretical models, and in addition provide a safe setting to test theories before taking them to the real world to make policy prescriptions. So if the goal of the enterprise is to advance our scientific understanding of political processes, laboratory experimentation should definitely be part of the picture.

I now make some remarks on the current status of the subfields of formal theory and modeling which may bear on the "split" between theory and empirical modeling:

One of the big problems in political theory, which distinguishes it from economic theory, is the absence of any general equilibrium theory from which to start the theoretical enterprise. The lack of general equilibrium in political science has put theoretical work in political science in a kind of limbo. Economists can frequently start from well accepted equilibrium models, and then do comparative statics by standard techniques. Political theory does not have well accepted equilibrium models to start from. So the theory must incorporate details of the situation. We agree that the underlying framework should be rational choice/game theoretic, but then to analyze a specific situation, details about the institutions must be part of the model. Thus, since the discovery of the generic non existence of equilibrium, the main trends in formal modeling have been to explicitly model the role of information, repetition and institutions in political processes. These trends have been accompanied by the increasing use of non-cooperative game theory, games of incomplete information, and explicit specification of extensive forms. Another trend has been to investigate evolutionary and agent based models. These take a fundamentally different view of behavior -- that individuals are programmed to behave in certain ways and only change their behavior through replacement or imitation.

Theorists like to prove very general results, but the institution based modeling goes in precisely the opposite direction, implying that there may be no general results. This opens up a big role for so called "applied theory." The implication is that empirical modelers, when studying a specific set of data, may not be able to find a model that they can just plug in and use. Rather, they may have to develop some of the theory themselves. This lack of ready made, plug in theories that empirical modelers can use in their work probably helps to explain some of the split between theory and empirical work. But it also opens up an opportunity for those researchers who can "speak both languages."

In summary, I think that the so called split between theory and empirical modeling is something that is natural, although perhaps greater than we would like because of the unique problems of social science. I also believe there are natural tendencies for researchers to bridge this gap, as it is in the interests of theorists to find empirical settings to test their theories, and in the interest of empirical researchers to have their research contribute to a deeper understanding of politics. The problem is that these natural correcting tendencies operate more slowly than we would like.

I do think that NSF can help to speed along the above process, and I think it would be helpful to the profession in the long run if it were

to play a role in this. I believe that NSF could clearly play a role in trying to populate the part of the profession that is conversant with both camps. I believe that the most effective ways of doing this are approaches that reach graduate students junior faculty, since they are at a stage of their career when they can still easily change their research focus or methodology. 1) One approach would be to help fund summer programs for graduate students, taught by faculty from both camps. 2) Also, I believe that a very promising strategy here is to fund small conferences which bring together theorists and empirical scholars. It is important that these conferences should include a large contingent of young scholars. There are already some conferences that operate on this mode (for example the Wallis Conference on political economy) and my impression is that they have been quite successful. 3) In keeping with my remarks above, I also think that funding of laboratory experimental research is very promising, since there is already a close connection between this kind of research and the corresponding theoretical modeling effort.
4) Finally, I believe that NSF could make a major contribution in bridging the divide by funding some major efforts to collect and assemble data that would be relevant to formal models. For example it would be a tremendous resource to develop a good time series of data which breaks down the receipts and outlays of the federal budget (in various categories) by congressional constituency, and also to try and estimate the incidence of new legislation by constituency.

Richard D. McKelvey
California Institute of Technology
Mail code 228-77, Pasadena, CA 91125

(626) 395-4091
(626) 793-8580 (fax)
rdm@hss.caltech.edu

Comments by Rebecca Morton

I thought about the history that led to the split a lot in writing Methods and Models and talk about it in chapter 1. First I would like to say that to some extent a split is necessary. That is, there will and should always be a place for what we think of as "pure theory" that is not intended or useful for direct real empirical analysis. The applied theory that we work with in empirical analysis will die if we try to kill off pure theory. There are some empirically oriented political scientists who argue that no paper should be published in a major journal without some empirical analysis in it. It is important that this group make it clear that such a stance would hurt the discipline (trying to go a tree without roots).

That said, the split in political science seems to me (and apparently others as well) bigger than it should be. Why? I think that are several reasons:

a. The discipline became more sophisticated in data analysis and this type of work became modal (the behavioral revolution) first and formal theory has been playing "catch up." In the process, as Dina points out, most political science began to believe that statistical models were theoretical models and that just writing a regression equation (and a some paragraphs justifying the form of the equation, or the methods of statistics, or the variables included or discarded) was theorizing. However, this is changing. In the paper with Chuck we sent out a request to many political scientists in an attempt to census the extent of formal/empirical research in political science. We were amazed by the quantity of our responses (over 200). We also surveyed the last five years of many of the major journals. Again, there is an increasing amount of formal/empirical work. That said, the majority (it varies by field) of the work we saw was the type of empirical analysis that is at best only loosely related to theory and much that is hardly influenced by theory at all. The fact that empirical approaches are more prevalent and have a stronger foothold in the discipline means that to the extent that we train undergraduates at all about what graduate level political science is like (and we don't do much of that, something I will turn to again below), we teach them statistics and, occasionally some "light" formal theory.

b. The unhappiness with and misunderstandings about rational choice among political scientists trained from a psychological perspective is, in my opinion, one big reason why we see less formal/empirical combined work. Many political scientists have an outdated view of what rational choice and game theory are based on work 30 to 40 years old with little beyond that. There is very little knowledge of behavioral game theory (which combines work out of psychology with standard game theory) and the use of what I called in MM almost rational choice and non rational choice game and decision theoretic approaches. I have done informal surveys of entry level intro to research courses in graduate programs and students are still taught the version of rational choice used in Downs for example as if this was current -- largely because the instructor doesn't know better either. While it is a classic and there are wonderful tidbits in it -- many students, who would find the more modern work more attractive, end up turned off by this work. To some extent the rational choice based modelers in political science are at fault here as well -- many do not like the behavioral work or the less than rational choice stuff. They have been fighting a hard battle for years to make a place for the rational choice approach. Thus I think that some rational choice modelers in political science are more dogmatic than perhaps in economics (I'm not sure about this since I may seem the more behaviorally oriented economists since many of the economists I talk to are experimentalists who are most aware of the places where rational choice can fail).

c. The fact that our graduate students come with very little math in their backgrounds or exposure to methods or modeling when they arrive. A student with this background who wants sophisticated training in both areas then must necessarily take longer to get his or her degree than he or she ever planned. Generally they take the sequence of methods classes. In the research design class they may all have to take they get taught some really old stuff as examples of formal modeling (mainly because they, and possibly the instructor, can't do the math of the better more modern stuff). They may still go on and take a one semester formal theory course, but rarely are they going to take much more. As a result we try to teach game theory and other formal approaches too quickly (in one semester) to students too ready to use it in their work (i.e. already close to dissertation stage). A one semester watered down game theory course for these students is not going to enable them to really understand what they are doing when they use it or know much if anything about the more sophisticated approaches in behavioral game theory, etc. Much less really know how to then take this work and apply it empirically.

On American Politics:

In our survey Chuck and I found that of the subfields American is doing the best in terms of this type of work. Yet there are important variations within American and problems in the subfield. The most applied theoretical headway in American is in the study of Congress and other legislatures. Yet, because Congress is a single unit, testing many of the theories that have been developed require a lot of ingenuity since the things that the applied theorists are often trying to explain don't vary (except sometimes over time, and then along with so many other things that the empirical analysis often doesn't tell us much)! What this shows is that the way we divide the subfields is obsolete. By dividing the subfields into American and other places impedes the kind of real empirical analysis of the theories that we develop -- the theories tend to be too institution specific and the data tends to be too dominated by American data. To some extent we see that researchers are trying to rectify this -- as in the book manuscript of John Huber and Charles Shipan where they take a theory developed about Congressional delegation and test it uses data on state legislatures and legislatures in other countries. But we need to drop American and Comparative as subfields and just have two main divisions, domestic and international and within domestic, legislatures, executive, etc. This is way overdue!!!!

The part of American that I think has the least amount of formal/empirical research is in the study of voting and elections. I think the problem here is the rational choice problem I mention above. Political scientists have a lot of hubris when it comes to examining voters -- they just can't believe that they might be as smart (or maybe more so) than academics! We like our elite models of politics primarily I think because we think we are the elite. Here I think more exposure to behavioral game theory might be useful as the researchers might find those approaches more acceptable given their biases about voters (and there is some of this going on -- the recent paper by Bendor and Diermeier is an example where you can't find dumber voters -- but it is a purely theoretical work in its current version).

Ways to bridge the gap:

a. Better undergraduate training. We need to teach our undergraduates more so that they can come to graduate school better prepared. The difficulty with doing this is that many public schools departments are evaluated based on undergraduate enrollments and most fear that doing so will cut these precious enrollments. Moreover, many students take political science because they falsely think it is useful for law school and we don't want to disabuse them of this notion. Fortunately some noted law schools are adding methods type classes and training, so we might be able to argue that we are increasing their potential at law school. We need to recognize that we may actually attract undergraduates if we teach them more about doing political science rather than gossiping about politics. I think that revamping undergraduate training in political science is the only way we will solve the problem in the long term. To some extent this is happening at a piecemeal rate, but there is still much more to be done.

b. In the short run, we need to provide some help to graduate students and encouragement for them to take more than one course in formal modeling but a sequence as well as a sequence in methods and courses that deal with combining the two. We need ways to fund an additional year or two of graduate school for students to complete both methods and modeling as well as their other coursework. Perhaps some centers could be established where students in programs that can't offer this training (either due to lack of faculty or students interested) can go for a year on an exchange program. A summer training program is not sufficient. Some programs already offer this type of thing (visiting for a year) but doing it is often difficult to manage if not impossible. For example, I had an _excellent_ student at Iowa who I made extraordinary efforts to get a chance to spend a year at Northwestern MEDS for this purpose. However, the deal fell through at the last minute when we had trouble convincing Northwestern that her math skills from Iowa were sufficient. In order to prove this, they wanted her to take some more math courses at Iowa and did not tell us until it was too late to do it that semester. In investigating taking the math classes at Iowa, the student became convinced that through the applied math program there she could take many of the same classes at Iowa as at MEDS and the hoops required to jump through for Northwestern were too difficult. The lesson here is that if the program is an away for a year thing we will no doubt lose a number of good students and that there may be a lot of ways of combining with applied math departments to provide such training.

The Formal and Empirical Divide In Political Science, With Comparisons to Economics and Epidemiology and Suggestions for NSF Remedies

Carl Simon
Professor of Mathematics, Economics and Public Policy
The University of Michigan
Director of the UM Center for the Study of Complex Systems

Background

For many of us the ideal social science paper includes both formal theory and empirical modeling. In these situations, there is a real phenomenon that drives and motivates the theory. The researcher chooses variables that are important for studying the phenomenon and posits relationships or properties regarding these variables. S/he then uses mathematical (or computational) analysis to draw conclusions from these relationships. Finally, the researcher uses a careful statistical analysis of real world data to test whether these conclusions hold in the real world.

By formal modeling, I believe that we mean either deductive formal modeling, in which a theorem/proof presentation is followed, or computational modeling, in which assumptions and equations are set forth, but the analytical problem is so complicated that computer simulation is used the draw conclusions from the model.

By empirical modeling, I believe that we mean a method by which data about some phenomenon is studied using only statistical tools. Ideally, there are some structural equations behind the choice of statistical method, guiding the study of how the variables are related to each other, but the empirical modeler does not manipulate these (usually linear or log-linear) equations to draw conclusions about their implications.

Comparing Political Science and Economics

Let me begin by comparing the prevalence of formal versus empirical modeling in economics (especially, microeconomics) and political science, although this background is probably well known to all the participants in this workshop. Microeconomics, especially as taught in upper-class and graduate microeconomics courses is a very mathematical subject. Its presentation usually includes assumptions, theorems and proofs. Naturally, this carries over into microeconomics research papers. In a survey of articles in two issues of the 1981 American Economic Review (AER), I found that 74% of the articles involved formal theory with no empirical work, while only 15% were empirical without formal theory. Only 11% involved both formal and empirical modeling,

By contrast, I surveyed the eighteen articles in two issues of the 1981 American Political Science Review (APSR). One of these eighteen centered on formal modeling (a piece by Bob Axelrod on the prisoner's dilemma). The other seventeen were purely empirical; none of the eighteen articles combined theory and empirics. In the early 1980s, I found by experience just how hard it was to get a theoretical piece published. Bill Keech and I wrote two theoretical articles on the optimal presidential term length. We were turned down by APSR precisely for being too theoretical for their audience. We ended up publishing one article in a book of papers that straddled the economics/political science boundary and the other in the Journal of Economic Behavior and Organization (where I was on the editorial board).

This situation has changed a bit in the last twenty years. It is now easier to get formal theoretical papers into the APSR, especially if the formal modeling involves game theory. In the sixteen articles in two issues of the 1998 APSR, I found that four articles were theoretical. The other twelve were purely empirical. Once again there were no articles that combined formal and empirical modeling. Hence, the need for this NSF workshop.

Interestingly, while the 1981 AERs were dominated by formal modeling; the 2001 AERs were dominated by empirical modeling. Of the 46 articles I examined in two issues of the 2000 and 2001 AER, 14 included only formal modeling, 21 only empirical modeling, and 11 included both formal and empirical modeling. In all, 70% of the articles included empirical modeling.

The statements by Dina Zinnes and Rebecca Morton in this collection describe a more extensive investigation of the economics and political science literature. Their conclusions are very similar to mine above.

In the 1981 AERs I examined, 85% of the articles included formal modeling; in 2001 the corresponding percentage was 54%. Economics seems much more rooted in formal models, while political science is rooted in empirical models. There are a number of circumstances that appear to reinforce these tendencies.

1. Microeconomic theory is taught as an analytic science in a theorem/proof mode, certainly to graduate economics students, and usually to undergraduate upper-classmen.
2. Political science is taught more informally at all levels, with only rare attempts at a theorem/proof format.
3. Economics undergraduate majors are usually required to take a rigorous calculus course. Economics graduate students are required to take at least one semester of mathematics for economists (usually taught from the book that I wrote with Larry Blume). Political science undergraduates and graduates rarely have such requirements.
4. Economics deals with rather concrete variables, like price, quantity, unemployment rate, that are fairly precise, relatively easy to measure and more conducive to rigorous mathematical analysis.
5. Political science deals with some concrete variables, like votes, but more often with harder to measure terms, like power, influence and affiliation.
6. These tendencies become solidified as older faculty members bring their experiences and prejudices to bear as they referee younger colleagues' research articles and grant submissions.

Economists are more often criticized for being too mathematical than political scientists are criticized for being too empirical. Possibly in response to or in anticipation of such criticism, economists have begun to blend their theory with empirical foundations and estimations.

How these two fields relate to game theory is a good basis for comparison. Most economics papers on game theory deal with developing the foundations of game theory with little regard to real world economics applications. On the other hand, political scientists often use game theory to understand real world political conflicts and conflicts-of-interest. Game theory, and its electoral special cases, has become the main analytical tool in formal political science theory. Political science graduate students are often required to take a course in game theory, unlike economics graduate students.

Comparisons with Epidemiology

Let me add one more comparison using an area in which I have written extensively: epidemiology. Despite its close relationship to biology, epidemiology is much more comfortable with empirical modeling than with formal modeling. The primary epidemiology journal, The American Journal of Public Health (AJPH), will generally turn down articles that have substantial formal modeling (including one I submitted last month). The major funding organization for epidemiological research, NIH, tends to support very few theoretical enterprises. Every five years or so, NIH sponsors a conference on "how to support more theoretical epidemiological modeling;" the end result is usually one RFP dedicated to formal modeling. However, since the review panelists are usually classical biostatisticians, funding for formal modeling remains difficult, at best.

Epidemiology's strong reliance on biostatistics leads to an interesting contradiction. The classical biostatistical "individual risk" approach involves – in its simplest form – writing out a large table with each row denoting a susceptible individual and each column denoting a "risky behavior." One notes which individuals with which risky behaviors became infected with the disease under study. One then performs a statistical analysis on this data and concludes, for example, that high cholesterol leads to a higher probability of a heart attack.

This individual risk approach has some serious drawbacks. It suggests correlations, but cause and effect is hard to determine. It is very linear so that the whole is strictly the sum of its parts. Finally, the underlying statistics requires strong independence assumptions. It is hard to swallow, especially in the study of contagious diseases, that person A's infection is completely independent of person B's infection.

I think that, just as in economics and political science, the best work in epidemiology mixes formal modeling with empirical data analysis. I am especially proud of my own work in this vein on HIV. My research group spent six years constructing and analyzing formal models of the spread of HIV until we were comfortable with all aspects of the models. Then, we compared the output of our model to empirical HIV data to estimate a critical parameter in the model – the contagiousness of HIV in a single sexual contact between an infected and an uninfected individual.

The Complex Systems Approach

Let me include one final observation comparing formal modeling in economics and political science. Economic theorists who analyze formal models of economic phenomena usually need to make strong simplifying assumptions to make their model mathematically tractable. Such simplifying assumptions often include: no imperfect rationality, no diversity, no dynamics, no learning or adaptation, and no organizational structure. Including these complexities is the hallmark of the complex systems approach to social science modeling. Being much more comfortable with a rigorous proof approach than with computer simulations, economists have been reluctant to include complex systems approach in their mainstream economics research.

On the other hand, political scientist modelers have been more comfortable with complex systems modeling and computational modeling in general. They are not so strongly tied to the reliance on formal mathematical proofs. They often care more about the very complexities that economists avoid, such as, learning, diversity, and organization structure.

Interdisciplinary Avenues

Since economists are more comfortable with formal modeling and political scientists with empirical modeling, one solution is to bring the two groups together in research teams on social science issues that call for both approaches. Institutions that have a history of making such interactions natural processes have true advantages here. Such institutions include:
1. Cal Tech, whose social science department includes both economists and political scientists; though formal models seem to be encouraged more than empirical ones.
2. Carnegie Mellon, whose Department of Social and Decision Sciences, includes both economists and political scientists, many of whom are adept at combining mathematical analysis, computational analysis, and empirical analysis.
3. Non-academic Institutions, like Rand, Brookings, and The Santa Fe Institute (SFI), where scholars from different disciplines or with different approaches from the same discipline naturally mix throughout the day in formal and informal research activities.

Please allow me to toot my own horn and describe why The University of Michigan (UM) should be included in this group.

1. The Institute for Social Research (ISR) at UM is the world headquarters of social science data collection and analysis. Under the leadership of David Featherman, it is striving to include formal modeling as the backbone of many of its empirical projects.

2. UM has one of the strongest group of formal modelers in its political science department, including Chris Achen, Bob Axelrod, Jenna Bednar, Michael Cohen, Ken Kollman, Skip Lupia, James Morrow, and Scott Page.
3. UM is the only academic institution with a center that is devoted to the complex systems approach to studying the sciences and that includes the whole range of the sciences: biological, physical, decision, engineering, and social. In fact, social science is a particular strength of the UM Center for the Study of Complex Systems (CSCS). Axelrod, Cohen, Kollman and Page play an active role in the activities of CSCS; Page is its Associate Director (and first hire). CSCS has recently applied for an NSF-IGERT (joint with SFI) that would foster complex systems approach to studying economic and political institutions. Its (unsuccessful) NSF Biocomplexity proposal had twelve investigators from twelve departments working together to understand how the underlying contact structure affects biological and social phenomena. CSCS and the Physics Department has recently hired SFI post-doc Mark Newman, who writes extensively on the role of networks in social and physical phenomena. UM graduate students who supplement their graduate work with five selected courses involving the complex systems approach can earn a certificate in complex systems.

Places like Cal Tech, CMU, UM, Brookings, and SFI would be natural catalysts for combining formal and empirical modeling. These are the kinds of places that encourage “thinking outside the box” and intermingling approaches from different fields. It is hard to catalyze work that combines different approaches in institutions with impenetrable walls between the disciplines.

Possible NSF Strategies

The NSF would do well to support institutes in such places where:

a) Formal modelers with an interest in the empirical aspects of their work can interact with empirical modelers who want to learn the basics of formal modeling. Bill Keech’s 1977-78 sabbatical at UM’s Institute for Public Policy Studies is an excellent example of how an empiricist who wants to learn more theory can do so successfully in the right atmosphere.
b) Such an institute can invite long term visitors who will interact over a semester or two and short-term visitors who will give a lecture or participate in a workshop. The University of Minnesota’s NSF-sponsored Institute for Mathematics and Its Applications (IMA) is a good example of a successful such institute; in that case the need was to bring pure and applied mathematical researchers together, as well as academic and non-academic researchers.
c) Such a place would be an excellent place for political science post-docs, who will be surrounded by more experienced social scientists with valuable experience in combining formal and empirical methods.

NSF can offer initiatives for individuals or teams of individuals, who can convince NSF of their interests and potential for combining theoretical and empirical modeling. NSF would need to ensure that the reviewers and panels for such proposals are amenable to the kinds of issues under discussion here. Some of such funding could be set aside for post-doctoral fellowships so that select new political science PhDs can have the opportunity to learn from individuals and groups that successfully combine the formal and the empirical.

NSF can support more courses on mathematical analysis and mathematical modeling in political science departments. In this vein, it can put forth an IGERT-like competition in which political science departments can gain more NSF graduate fellowships if (nearly) every new PhD thesis with formal or empirical modeling content has both a formal and an empirical modeling component, or in which the appropriate faculty demonstrate an increased willingness to include both kinds of modeling in their published papers.

Using the simple analysis of the formal and empirical content of journal articles as I described earlier, the NSF can keep track of the success of this initiative, either in the science as a whole or in the work of individual researchers. In the long run, the hope is that the percentage of published articles that include both formal and empirical modeling in the major political science journals will increase as a result of the initiatives spawned by this workshop.

(I would like to thank Bob Axelrod for a fruitful luncheon conversation about some of the issues raised in this paper.)

Comments by Peyton Young

Our topic is the supposed "split" in political science between formal theory and empirical modeling, and how NSF might help to bridge it. Here are my preliminary thoughts on the subject, which (not surprisingly) overlap comments by some of the other participants.

The problem as stated is certainly not unique to political science. For example, the split between formal theory and applied econometrics is, in some ways, even worse in economics. The reason is not hard to find: the different branches of formal economic theory (macro and micro) and econometrics (time series and applied micro) each require years of specialized training. Individuals specialize in theory or econometric/empirical work according to their comparative advantages. The "split" is therefore due to the increasing specialization within disciplines. This is unfortunate, but it will be difficult to arrest, and it is certainly not peculiar to political science.

There is, however, a different way of framing the problem that is more pertinent to political science. The question is whether empirical work *complements* formal theory in the sense that theories about political behavior and institutions are routinely tested empirically and theory guides the relationships that empiricists try to test. Here there certainly is a difficulty as compared with economics. The essence of the problem is that political science lacks an accepted, parsimonious theory that convincingly relates individual behavior with the functioning and form of political institutions. (I am not among those who believe the rational actor model provides such a theory, although it seems to be the best candidate at the moment, and it certainly has had some partial successes.) I think it is fair to say, however, that a good deal of empirical work in political science proceeds without benefit of theory, and a good deal of theorizing goes on without benefit of empirical validation -- in some cases with no possibility of validation.

Of course, one sees the same thing in other subjects, including economics, but by and large empirical economics tests propositions that flow directly from theory. And admittedly there is a lot of esoteric theory being spun by game theorists (and the remains of the general equilibrium crowd) that probably will never be tested against data. Yet there is a growing sense within economics that theory needs to be reworked substantially in order to come to grips with economic phenomena. I am not at all sure that the same kind of searching criticism and reformulation of formal theory is going on within political science. Of course, there are quite a few people who never bought into it in the first place, but among the formal theorists it seems that enchantment with the rational actor model has never been higher. This is strange, since within economics the foundations of the rational actor model (perfect foresight, hyper-rationality, common knowledge, standard discounting, expected utility maximization) are being called into question. Indeed, the thrust of recent work in behavioral economics is that people behave in ways that are *boundedly rational*, and that their motivations are far more complex than standard theory has assumed.

Complementing these empirical developments, theory is starting to focus on the implications of bounded rationality, limited foresight, and learning for the dynamical behavior of aggregates of individuals and the evolution of institutions. In other words, the new questions in economics are about comparative dynamics rather than comparative statics, and about the complex motivations (rational or otherwise) that guide individuals in making choices. These same issues would seem to be at least as pertinent for political science.

To return to the topic of the workshop: the task as I see it is not how to bridge the gap between formal theory and empirical modeling, but to make them more complementary and bring them to bear on the same questions. Here political science is in some ways at an advantage compared to economics. Traditionally, political science has entertained more of a heterodox approach to explaining political phenomena; there are more competing points of view and a healthy respect for empirical evidence. At the same time, disputes rage about which approaches (formal or otherwise) are most appropriate, yet these debates are largely sterile if they only occur at a philosophical level.

Political science is at a juncture where more diverse interpretations of "formal theory" can and should be developed and tested. On the one hand this means opening up rational actor models to some of the new developments in bounded rationality, learning, and evolutionary modeling. On the other hand, it means exploring alternatives to rational actor models that are formal (in the sense of logically coherent and mathematical) but tailored to the analysis of particular institutions. Early voting theory (in the sense of Black and Arrow) is one example of this kind of theory -- while it has strategic aspects, it is essentially a separate, normative theory of collective decision-making. Similarly, theories of fair division provide a formal framework for analyzing certain aspects of political decision making and the structure of political institutions, but they are not the same as rational actor models. I am sure there are other examples.

How can NSF help promote a better integration of formal theory and empirical modeling within political science? One way is to fund workshops in which different theoretical approaches (rational actor, evolutionary, agent-based) are brought to bear on the *same* problem. For example, how far do these different frameworks go toward explaining the following questions: Why do rates of political participation differ substantially across countries? When does "unrest" coalesce into "revolt"? Why do transitions to capitalism appear to succeed in some countries and fail in others? And so forth. The idea would be to invite representatives of different theoretical frameworks to explain a given phenomenon, and then invite empirical methodologists to design tests of these supposed explanations. Not only would this provide a forum for comparing the strengths and weaknesses of different conceptual frameworks, it would undoubtedly suggest new types of data that need to be collected in order to answer these questions more definitively.

Another way for NSF to help integrate theory and empirical work is to support interdisciplinary groups of researchers that combine theoretical and modeling expertise on the one hand with empirical and experimental expertise on the other. For example, in addition to political scientists, these research groups might include anthropologists, economists, experimental psychologists, and computer scientists, depending on the problem being studied. A similar approach has been adopted by the MacArthur Foundation, which sponsors interdisciplinary research networks in particular areas of economics such as inequality, growth, preference formation, etc. These have been quite successful, both in re-orienting the research of senior members of the profession and in exposing younger members (graduate students and post-docs) to new ways of thinking that have not yet entered the standard curriculum.

Formal models and statistics in IR: the whys and wherefores

By

Dina Zinnes
University of Illinois

Both modeling and statistics are very much alive and well in the field of international relations. I surveyed the major journals over the past ten years for an APSA roundtable, and discovered that both forms of research, together, now dominate the literature. That's the good news. The less than good news, from my perspective, is that statistically based research studies far exceed articles that use mathematical models and the number of articles that combine a mathematical modeling argument with empirical statistical analyses is almost non-existent. I also found that the far and away dominant mode of mathematical modeling is game theory. These facts are both surprising and not surprising.

I find them surprising for three reasons. First, although both statistics and mathematical models are forms of mathematics, the contribution that each makes to the research enterprise is very different. More importantly, they seriously need each other if the science of international politics is to progress. Furthermore, that statistical analyses should dominate the field is like having a cart without a horse. Statistics and empirical work should *follow* the careful development of an argument/theory; ideally, statistical analyses should come at the conclusion of the statement of a mathematical model and the determination of its deductions qua testable hypotheses. Theory, mathematical model, deductions, hypotheses, measurement and research design, data collection and statistics --- this is what I believe to be the appropriate progression. Finally, to discover that most mathematical models are game theoretic suggests that all questions in international politics involve overt decision making by rational actors. For those of us intrigued by systemic problems like the operation of the balance of power, this reliance on one paradigm seems to cheat us of the richness of the ideas of theold masters in the field.

On the other hand, the discovery that statistical research rules the world of international political research is not at all surprising if we look at the history of the field. IR moved from the legal and discursive realm into the behavioral world via empirical and statistical analyses. The folks that moved the field into a more analytic posture came largely from allied social sciences --- psychology primarily --- where the bent was empirical and the object was to measure, collect data and do statistical analyses. Although early in the behavioral movement JCR published Rapoport's overview of Lewis Fry Richardson's work (1957?), few were able to read and appreciate what Richardson had done. How many political scientists had any idea of what a differential equation was? Collecting data, on the other hand, while challenging when you had to use historical material rather than run experiments, was something that most could do. And, while somewhat alien, statistics was not that difficult to comprehend --- though plenty of mistakes were made in those early years. Add to this the coming of the infant computer age and statistical packages and there is little more that needs to be said.

These events produced a dynamic of their own. The statistical mistakes that were made lead to cries for better training in statistics. Initially political science programs solicited the aid of sociology and psychology to help teach graduate students statistics, but within a relatively short time frame it became obvious that the problems of allied fields are different from those in political science and so the statistics needed are not necessarily the same. Analysis of variance, for example, is ideal for experimental settings but somewhat questionable when used with historical data. So graduate programs begin to train their own students in statistics. As the quantitative tradition took hold and the two language requirement was replaced by one language + methods, e.g. statistics, and then no languages and a greater emphasis on data collection and statistics, graduate programs changed. Before too long it was argued that everyone needed statistics and methods and it became a requirement of the major graduate programs.

This evolution made for greater clarity in research programs. In international relations it forced investigators to be explicit and systematic about the questions they asked and their attempts to provide answers. But at the same time it moved the discipline from a more theoretical posture to a more descriptive one. If you couldn't measure it, collect data, set up a null hypothesis and run a statistical test the problem was of questionable relevance and significance. While the mushy arguments surrounding "theories" like realism, idealism, balance of power, etc. were not going anywhere very fast, the quantitative approach almost threw the baby out with the bath water. Potentially interesting arguments, albeit fuzzy in their construction, were reduced to one-line regression equations, as training programs produced students, soon-to-be researchers, with the regression mind set. Data collection methodologies and statistics, which normally follow a theoretical argument, became, instead, the driving force behind the arguments.

So the behavioral revolution with its emphasis on quantitative methods, together with graduate training programs that reinforced this development, set the stage for the current research environment. But two additional factors contribute to making the statistical paradigm preeminent. The first is the confusion that surrounded the differences between statistics and mathematical modeling. For example, Hayward Alker's Mathematics and Politics describes statistical analyses, indices and mathematical models almost in the same breath leaving the reader with the impression that they are all the same thing. It was assumed that statistical models were mathematical models. Which indeed they were. But collapsing the theoretical argument into the statistical analysis reduced potentially exciting ideas to relatively mundane and simplistic statements.

The other factor that played a role in minimizing the importance of mathematical modeling was the inadequacy of the undergraduate background of the typical political science graduate student. Most students that enter graduate programs in political science have almost no background in mathematics. Discovering towards the end of the first year of a graduate program that mathematical modeling is more than, and indeed different from, the two semesters they have devoted to learning statistics, puts the typical student in a quandry. There is no longer sufficient time to take the multitude of mathematics courses needed to become a proficient modeler. If you don't have a mathematical background its too late to get it. Or, you move into game theory or computational modeling.

While game theory and computational modeling can become highly mathematical it is feasible to cast problems in either of these media with minimal training. Instead of half a dozen mathematics courses, a course or two in either of these venues will equip the researcher with sufficient background to permit a quasi-modeling research program. And of the two, game theory methodology is the most accessible in the typical university environment.

The above chronology helps to explain not only the dominance of statistics but also the reason why there is so little research that combines mathematical models with statistical analyses: not only have statistical models usurped the role of mathematical models, but few researchers have background in both arenas. Moreover, because statistical models are used as mathematical models it is unclear how the two might be combined. How do you test a game theory model without stacking the cards in your favor? What data and statistical test provide one with evidence that a given set of differential equations adequately capture a particular problem? In short, making a mathematical model face the empirical facts is a new venture requiring new ways of thinking about theories, modeling, data and tests and raises intriguing questions about the testability of the deductions they produce.

From here to there: how can we change the current picture?

The above outline provides some clues as to how we might change the political science landscape and suggests a very significant role for the National Science Foundation in facilitating this dynamic. Clearly the bottom line is training. But to jump-start the discipline in this direction requires imagination and sensitivity to what is feasible. Graduate students and faculty can't get a PhD in mathematics. Indeed it's not clear that even if this were possible it would accomplish the objectives. Here are some possibilities:

1. Mathematics in short course form. The mathematics department at the University of Illinois has been engaged in developing "discovery" interactive teaching modules through the use of Mathematica. A number of modules in a variety of mathematical fields now exist. Careful choice of subsets of these packages could be made available for those needing background in algebra, calculus, and differential equations. While the more mathematics the better, the typical modeler does not need to prove every theorem in differential equations to understand the main ideas.

2. **Translation of theories into models.** It used to be surprising to me that those with extensive background in mathematics nevertheless stumble when they attempt to use their training to develop models of meaningful questions in a field. But over the years I have come to appreciate this difficulty. You may speak English well and you may speak Mathematics well, but translating the former into the latter is an art unto itself. We therefore need to develop materials that train potential modelers in this translation process. There are two components. First, researchers must have some idea as to the appropriate mathematical language for their particular problem. Mathematical languages say some things well and some things poorly. Sensitization to what can and cannot be done with various forms of mathematics is therefore essential. Second, once inside a chosen mathematics one needs to know how to make the appropriate mapping.

3. **Fitting statistics to models.** Political scientists, mathematicians and statisticians need to combine forces and think through the new issues raised by mathematical models. Keeping in mind that the goal of testing is to gain confidence in one's ideas, we must develop procedures that make it possible to reject models. Only if we know what it would take to reject a model can we gain some confidence that our theory/model says something about the real world. Some standard statistical

tests will be appropriate, but in many cases very different protocols will be needed.

The National Science Foundation can make these things happen in many ways. Through appropriate funding it can facilitate the composition of the needed mathematical course materials and make them available on web sites to provide avenues for more mathematical training. It can support research teams of mathematicians and political scientists to develop translation tools that can aid researchers in understanding how to move their problem into appropriate mathematics and sponsor short workshops to teach the translation process. It can hold conferences and commission papers that address the philosophy and mechanics of testing types of mathematical models.

APPENDIX C

The Empirical Implications of Theoretical Models: A Proposed Workshop for Political Science

By

Jim Granato and Frank Scioli[17]

Synopsis

Objective: Improve the technical proficiency of political science.

Timing and Venue of the Workshop: July, 2001 at the National Science Foundation.

Participants: Senior scholars who have a proficiency in various technical areas or who have actively participated in activities that improved the technical expertise of the discipline or both.

I. The Issues

The past two decades have witnessed an enormous improvement in the technical competency of scholars in political science. Progress, however, has been uneven. A significant concern is that a schism has developed between those who do formal (theoretical) work that is highly mathematical and those who do empirical work that tends to emphasize applied statistics.

This divide begins in graduate school in political science where technically oriented students take the requisite courses in formal modelling or statistical methods, but then go on to improve their skill level in only one or the other of these fields. The reason for this type of technical specialization is most likely a function of the time constraints faced by students in graduate school. A typical four to five year Ph.D. education in political science does not allow enough time for a graduate student to become comfortable and competent in both technical areas.[18]

---

[17] We would like to thank Bill Butz and Cheryl Eavey for their comments and suggestions.

[18] Students invariably must take courses in their substantive subfield (if they are not methods/formal majors). In the case of fields such as comparative they must engage in field work or learn a foreign language or both.

The end result is that a good deal of research in political science is highly competent in one technical area (formal modeling or applied statistics), but lacking in another.

This divide has been harmful to the progress of science. The consequences take many forms. One important symptom[19] of this technical separation is a break in microlevel and macrolevel theorizing and analysis, often leading to contradictory and incoherent predictions when one moves from one level of analysis to the other. Formal work tends to view things from an individual level and empirical research deals with aggregate analysis, either summing over individuals, goods, or events. In the book, Cross Level Inference (1995), two political scientists, Chris Achen and Phil Shively, argue for a union of the two levels of analysis and justify it accordingly:

> ...we follow Green (1964) in preferring, when the microlevel is the theoretically meaningful domain, that macrolevel models be explicitly derived from micro-level assumptions. Without that constraint, macrolevel research too easily slips into studies of the interrelationships of meaningless statistical aggregates. Only when both macrotheoretical propositions and statistical assumptions are rigorously inferred from the microlevel can we have faith in the macrolevel studies (pp. 25).

The real world consequences of this schism have been substantial as well. For example, a wonderful real world experiment of macroeconomic theory and practice informs us on the pitfalls of failing to adopt the Achen-Shively research strategy.[20]

This concerns unemployment in the late 1960s and 1970s and the attendant policy/theory brought to bear on it. This is sometimes referred to as the breakdown of the Phillips curve. To be concise, the breakdown of this "theory" was due to a failure to reconcile microlevel theory with macrolevel outcomes (Friedman, *American Economic Review*, 1968; Phelps, *Journal of Political Economy*, 1968). There were four important consequences of this breakdown:

1) policy mistakes caused harm (in varying degrees) to millions of individuals in the 1970s in terms of lost employment opportunities;

2) there was a shift away from consideration of individual welfare losses;

---

[19] This is not meant to limit discussion to just the issue of microlevel and macrolevel issues, it is only an illustration of what has occurred because of the technical divide.

[20] We would like to use a political science example, but our discipline has not matured scientifically --- at least not as economics has. By *scientific maturity* we refer to the technical nature of substantive debate that spans subfields. In political science, for example, *there is* work on *ecological regression* (Achen and Shively, 1995; King, 1997) but that work is considered more the province of political methodologists and not students of the various political science subfields. Indeed, it is not part of a broad theoretical debate *within* a given subfield. There are individual (read isolated) exceptions by various authors in adopting a micro- and macrolevel analysis (see Brady (*Political Analysis*, 1993) and the citations therein).

3) there was a shift in focus away from microlevel reasoning on job search and job matching which would have informed policymakers about unemployment duration, earnings distributions, and other issues that focus on the well-being of labor; and

4) there were incorrect conditional forecasts.

II. The Opportunity

A linkage between micro and macro level analysis --- one type of linkage between formal and empirical approaches --- presents a vast array of opportunities for the advancement of research in political science. These opportunities include analysis of social science problems that deal (among other things) with: a) multiple goals of citizens (with and without limited choices); b) the endogeneity of rules; and c) preference changes (including regime shifts).

This linkage also requires aggregation over individuals.[21] This presents a further opportunity to explore (the problems outlined above) and extend aggregation theory from the perspective that may have escaped economists and students of aggregation theory.

Aggregating over citizens has a long history.[22] For example, in the study of aggregate demand functions there is notable work by Debreu (*Journal of Mathematical Economics*, 1974), McFadden, Mas-Colell, Mantel, and Richter *(Journal of Economic Theory*, 1974), Sonnenschein (*Western Economic Journal*, 1973), and Mantel (*Journal of Economic Theory*, 1974 and 1976). They demonstrate that there is little about which we are sure. Among these papers some general patterns emerge including, a lack of integrability and restrictive homogeneity assumptions (see Sonnenschein, 1973).[23]

In this particular case --- aggregating over demand functions --- the source of the problem economists face is the heterogeneity in citizen preferences (tastes). However, from a political science perspective, as indicated by issues such as multiple goals and the like (outlined above), it could be possible that differences in preferences (tastes) are a non-issue. As such, it may be the nature of political science problems and research questions that may lead to ways (opportunities) to circumvent aggregation problems that trouble economics.[24]

[21] While the focus here is on individual citizens, it is also possible to aggregate over individual nation-states or other units of analysis.
[22] The characterization that we are describing applies to the "best" research in economics. As one moves below this the terrain becomes murkier and the field becomes suffused with the same weaknesses in analysis (research design) as political science.
[23] Forni (1998) is a recent reference on the aggregation problem in macroeconomics.
[24] To capture variations in tastes and other variables of interest requires knowledge of the joint distribution of all the taste determining variables and other variables of interest. Unfortunately, as stated above, this is difficult to specify, but that is what political scientists need to be made aware of. It is also an exceptional opportunity to advance social scientific inquiry. And it is the ultimate payoff where research not only centers on important substantive questions, but due consideration is given to how to analyze the problem and identify the parameters (read behavior) of the model.

III. The Role for the NSF

The NSF can play a role in attending to this issue. A workshop is proposed to facilitate discussion among leading scholars who share these concerns and who may have some suggestions on how to correct the problem.

Among the issues to be considered by workshop participants:

1) Identification of the role that NSF can play through funding opportunities and program initiatives in advancing the linkage of formal modeling and applied statistics.

2) Identification of a coherent strategy for policy implementation across disciplines and whether the various initiatives should be instituted at the same time or sequentially.

3) Identification of scholars who can revisit this issue on a periodic basis to determine progress in political science.

IV. Invited Participants[25]

The original list of participants includes:

1. Christopher Achen (University of Michigan)
   (E-mail: achen@umich.edu)

2. John Aldrich (Duke University)
   (E-mail: aldrich@acpub.duke.edu)

3. James Alt (Harvard University)
   (E-mail: jalt@latte.harvard.edu)

4. Henry Brady (University of California, Berkeley)
   (E-mail: hbrady@csm.berkeley.edu)

5. Randall Calvert (Washington University, St. Louis)
   (E-mail: calvert@artsci.wustl.edu)

6. John Freeman (University of Minnesota)
   (E-mail: freeman@polisci.umn.edu)

---

[25] Attention has been paid to issues regarding the participation of women and minorities. Given the complexity of the topic, necessary expertise (and seniority) instructs the list of invitees.

7. Carol Graham (Brookings Institution)
(E-mail: cgraham@brook.edu)

8. William Keech (Carnegie Mellon University)
(E-mail: keech@andrew.cmu.edu)

9. Richard McKelvey (California Institute of Technology)
(E-mail: rdm@hss.caltech.edu)

10. Rebecca Morton (University of Houston and New York University)
(E-mail: rmorton@uh.edu)

11. Carl Simon (University of Michigan)
(E-mail: cpsimon@umich.edu)

12. H. Peyton Young (Johns Hopkins University)
(E-mail: pyoung@brook.edu)

13. Dina Zinnes (University of Illinois, Urbana-Champaign)
(E-mail: d-zinnes@uiuc.edu)

APPENDIX D

EITM Workshop Agenda

Monday, July 9, 8:30 a.m.-6:00 p.m., Room 1235

8:30-9:00 a.m.: Introductions and Preliminary Considerations

Opening Statements from Norman Bradburn and William Butz

9:00-10:30 a.m.: Discussion Point 1: Identification of the factors contributing to the split between formal theory and empirical modeling (*All).*

Current Status of Subfields:

a) American government and politics (*Aldrich, Keech, and Morton as discussion leaders*).
b) Comparative government and politics (*Alt and Freeman as discussion leaders*).
c) International relations and politics (*Zinnes as discussion leader*).
d) Methodology, Modeling and Statistics (*Achen, Brady, and McKelvey, as discussion leaders*).

Current Status of Other Disciplines: (*Simon and Young as discussion leaders).*

10:30-10:45 a.m.: Break

10:45-12:00 p.m.: Continuation of Discussion Point 1.

12:00-1:00 p.m.: Lunch

1:00-2:45 p.m.: Discussion Point 2: Discussions of the need (and feasibility) to bridge formal theory and empirical modeling and of viable strategies for doing so in the discipline (*All*).

2:45-3:00 p.m.: Break

3:00-4:15 p.m.: Discussion Point 3: Discussion of interdisciplinary avenues and extensions (*Eavey as discussion leader*).

a) Academic examples (Carnegie-Mellon, Cal Tech)
b) Institute examples (Brookings, Santa Fe)

4:15-4:30 p.m.: Break

4:30-6:00 p.m.: Discussion Point 4: Identification of the role that NSF can play through funding opportunities advancing the linkage of formal modeling and empirical modeling (*All)*.

a) What has proven effective in the past?
b) Are there best practices in other disciplines?

7:00 p.m.: Dinner at Tutto Bene (A conference room will be reserved and directions will be provided. It is a short walk from NSF.)

Tuesday, July 10, 8:30 a.m.-12:00 p.m., Room 1235

8:30-10:30 a.m.: Discussion Point 5: Identification of a coherent strategy for implementing the initiatives in the discipline via a "Dear Colleague" letter from NSF (*All*).

a) EITM: Dear Colleague letter (what form and what priorities)

  i) Infrastructure opportunities.
  vii) Annual meetings (how many, single subfield (problem) or open).
  viii) Graduate and/or undergraduate student opportunities.
  ix) Junior and senior faculty opportunities.
  x) Inter/multidisciplinary opportunities.
  xi) Comparative (joint) opportunities.
  xii) Other considerations.

10:30-10:45 a.m.: Break

10:45-11:45 a.m.: Discussion Point 6: Discussion of measurable indicators that indicate progress in the discipline (*All)*.

11:45-12:00 p.m.: Discussion Point 7: Identification of scholars (junior and senior) who can revisit this issue on a periodic basis to determine progress in Political Science (*All*).

12:00-12:30 p.m.: Departure from NSF

APPENDIX E

A Letter to Political Science Colleagues from The Political Science Program at the National Science Foundation, July 26, 2001

Dear Colleague:

This letter provides information about three opportunities in the Political Science Program at the National Science Foundation. These opportunities are part of a continuing investment in improving the technical proficiency of Political Science by bridging the divide between formal and empirical approaches.

I. The Empirical Implications of Theoretical Models (EITM)

To that end the Political Science Program convened the Empirical Implications of Theoretical Models (EITM) Workshop on July 9-10, 2001. In both written commentaries and contemporaneous statements, Workshop participants recommended that the Political Science Program address the technical divide in three broad areas:

- Education: Training and Retraining
- Dissemination of Knowledge: Conferences and Workshops
- Research: Establishment of Research Work Groups

The EITM Workshop recommendations, as well as the scientific and infrastructural progress and needs of Political Science, are the bases for funding the following priorities. The Political Science Program may fund up to $1,000,000 for these activities for fiscal year 2002. The specifics of each funding area follow. The Political Science Program stipulates at the outset that **all** proposals for **all** three priorities must adhere to the following guidelines.

It should be noted that the Political Science Program will be monitoring all EITM activities for evaluation purposes. Future EITM Workshops will be held to determine future budget allocations based on the success of the following priorities and to extend EITM into other ventures not covered at this time.

II. For All Opportunities

- All EITM proposals **must** contain a **formal and empirical component**. Any proposal submitted under the auspices of EITM that does not have this explicit linkage **will not be reviewed** in the EITM competition. Those proposals that do not comply will be processed under the normal competition instead. There are no exceptions.

- The formal component and empirical component must be explicitly outlined. Formal components include (but are not limited to) game theory and dynamic stochastic modeling. Empirical components include (but are not limited to) applied statistical procedures and experiments. “Hybrid” techniques such as agent-based modeling are also welcome.

- The Project Description section of the proposals must not exceed NSF's standard length of **15 pages**. In the Description, proposals must explicitly describe how their designs offer a combination of **formal *and* empirical** analysis. Principal Investigator(s) must discuss how their proposal advances progress in reducing the formal and empirical divide. Proposals further must include a dissemination strategy to ensure that the output of the activities reaches a large audience. **Appendices will not be authorized.** Proposals may refer to supplementary materials and other directly relevant information, posted on investigators' publicly available Web sites. Reviewers will be asked to safeguard their anonymity when accessing these Web sites.

- All EITM proposals must be submitted through NSF's FastLane. The target date for submission for activity ***IIIa*** and ***IIIb*** is **January 22, 2002** and for ***IIIc*** is **June 15, 2002.**

- The cover page for the proposal should identify itself as "**Political Science Program: EITM Competition *IIIa*, *IIIb*, or *IIIc*"**. For example, if the investigators' submit a proposal for competition ***IIIa***, then the cover page should read: **Political Science Program: EITM Competition IIIa.** This should be placed on the upper right side of the cover page on line two in the space titled **"For consideration by ----."** Line one of the upper right hand corner of the cover page should have the due date: **January 22, 2002.**

- Site visits involving the most meritorious proposals may be conducted, as appropriate.

- All EITM proposals must include a plan for efficient, effective, and responsible project management and for fair, inclusive, and open personnel selection.

- All EITM proposals must document that costs are commensurate with activities and objectives. A proposal also must discuss how its project will continue, if appropriate, when Political Science Program support ends. Proposals may include a plan for the generation of other matching funds.

- The Political Science Program expects that the applicant receive a significant financial and logistical commitment from their host institution to supplement the NSF (EITM) proposal.

- The Political Science Program encourages, when practicable, incorporating scholars and students from recognized and respected programs and institutions outside the United States in EITM activities.

- The Political Science Program encourages, when practicable, interdisciplinary linkages.

The EITM Opportunities

IIIa. Continuing investment in Political Science Education: Training and Retraining

- The Political Science Program may make **at least one award** to a meritorious proposal for intensive 4-week summer training institutes. The award will be used to introduce, train, and enhance both the formal and empirical skills of participants. The maximum allocated for this is $600,000 total for a four-year award. The expected start date and expiration date are July, 2002 and July, 2006.

- Scholar-investigators who possess the formal, empirical (or could team with others), and administrative skills, as well as the necessary resources, to undertake an important, large-scale, summer training institute are invited to submit proposals.

- Participants eligible to receive said training and retraining include graduate students, post-docs, untenured faculty, and tenured faculty.

- In the event there is more than one summer institute under operation, it is expected that linkages will be established between the various institutes to further the dissemination of knowledge to all participants and the scholarly community at large.

- Contingent on the outcome of the EITM Education/Retraining proposals and the availability of funds, the Political Science Program will conduct another competition in four years.

IIIb. Dissemination and Continuity of Formal and Empirical Linkages in Political Science: EITM Workshops and Seminars

- The Political Science Program may make **awards** to the most meritorious proposals for periodic workshops and seminars that promise to advance the dissemination of basic research that links formal and empirical analysis. The maximum allocated for this priority is $300,000. The annual amount of each award is expected to be $50,000 total; the duration of an award is one, two, three, or four years. The expected start date is July, 2002; the expected expiration date is not earlier than July, 2003 or later than July, 2006.

- Each individual workshop or seminar **must have a specific theme or problem** that allows for a variety of analyses which link formal and empirical approaches.

- Scholar-investigators who have the requisite skills and resources to administer the workshop and seminars are invited to submit proposals.

- Participants in these workshops and seminars may include a mix of graduate students, post-docs, untenured faculty, and tenured faculty.

- Where practicable workshop and seminar organizers are encouraged to establish linkages with the summer institutes and the possibility of organizing joint ventures.

- Contingent on the outcome of the EITM workshop and seminar opportunities, and the availability of funds, the Political Science Program may offer award opportunities on an annual basis.

IIIc. Research with Formal and Empirical Linkages in Political Science: EITM Research Work Groups

- The Political Science Program may make **awards** to the most meritorious proposals for planning workshops that foster the establishment of research work teams that advance the dissemination of basic research linking formal and empirical analysis. The maximum allocated for this priority is $100,000. The annual amount of each award is expected to be $20,000; the duration of an award is up to two years. The expected start date is July, 2002; the expected expiration date is not earlier than January 1, 2003 or later than January, 2005.

- Each individual research work group workshop **must have a specific theme or problem** that allows for a variety of analyses which link formal and empirical approaches.

- Scholar-investigators who have the requisite skills and resources to administer the planning workshops are invited to submit proposals.

- Participants in these planning workshops may include a mix of graduate students, post-docs, untenured faculty, and tenured faculty. That number shall not exceed 12 total members.

- Upon completion of the workshop, participants are eligible (as a team) to compete in the regular Political Science competition or future EITM research competitions.

- Contingent on the outcome of the EITM planning workshop opportunities, and the availability of funds, the Political Science Program may offer award opportunities on an annual basis.

Questions should be addressed to Dr. Frank Scioli or Dr. Jim Granato, Political Science Program Directors, National Science Foundation. E-Mail: fscioli@nsf.gov; jgranato@nsf.gov. Phone: (703) 292-8762.