CAMBRIDGE

Cambridge University Press
0521853621 - Computational and Mathematical Modeling in the Social Sciences
Scoot de Marchi
Excerpt
More information

1

Not All Fun and Games

Challenges in Mathematical Modeling

INTRODUCTION

In large part, the inspiration for this book came from three sources, which can be categorized neatly as a failure, a challenge, and an ideal. First, the failure. When I began teaching in the profession, I was immediately assigned to graduate methods coursework. This is the experience of many professors trained in the last decade with a mathematical bent, and I was lucky enough to teach at an institution with an excellent culture. Unlike many other political science departments that exist in a state in which "there is war of every one against every one," Duke's political science department is almost entirely free of disputes about the value of mathematical modeling in the social sciences. Divisions of opinion certainly exist but, more or less, everyone in the department recognizes the virtue of mathematical methods for at least some problems.

Better still, even those who do not practice mathematical modeling believe in good research design. As many prospective faculty members discover during their job talks, "methods questions" and questions about research design are just as likely to come from the theorists of the department as anyone else (though couched in different terminology). Between job talks, faculty brown bags, and informal interactions graduate students have with faculty, it would be hard to finish a Ph.D. at Duke and not try your hand at mathematical modeling.

Despite this positive culture, teaching graduate methods coursework has not been easy. As has been noted in numerous places, the

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521853621 - Computational and Mathematical Modeling in the Social Sciences
Scoot de Marchi
Excerpt
More information

shock most politics students experience on entering graduate school is severe. They expect to talk shop, debate the issues, and deal with "big" questions about the state of the world; instead, their first experience of graduate training at Duke involves a mathematics camp in the dull heat of August. No weighty matters of politics are discussed in this camp, unless one thinks that urns and the different colors of balls one places in them are of great import. Some students take years to get over this shock, essentially repeating much of their methods coursework when they come to a point in their own research where they have a pressing need for it. Others acquire good technical skills but nonetheless have great difficulty finding interesting questions or arriving at "good" models. Clearly, my best efforts were not sufficient and it drove me to think about issues of modeling in the social sciences and how one should attempt to improve matters.

In particular, why were so many bright graduate students, many of whom had good technical skills, unable to make the leap to generating testable theories? Why did many graduate students identify themselves primarily by their choice of method (e.g., game theory) rather than their research question? And, finally, were there any features of mathematical methodology in political science that added to the difficulty of training graduate students? These questions form a thread that continues throughout this book, and, hopefully, the questions offered here will demonstrate that many of the problems in training are related to conceptual problems in our mathematical methodologies.

The second influence on this book concerns a challenge to the discipline raised by Beck, King, and Zheng (hereafter, BKZ). Their paper – "Improving Quantitative Studies of International Conflict: A Conjecture" – appeared in the *American Political Science Review* in 2000. The paper was a broad challenge to empirical work throughout the social sciences, not just in international relations, and turned on the idea of what the proper relationship was between deductive models (usually represented by game theory) and empirical work (applied statistics). Normally, the ideal paper for the mathematical modeling crowd is a well-specified game that reaches some equilibrium outcome, which is then instantiated and tested in an appropriate statistical model. If hiring is any signal of departmental preferences, empirical work or game theoretic work alone is not as desirable as a combination of the two.

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521853621 - Computational and Mathematical Modeling in the Social Sciences
Scoot de Marchi
Excerpt
More information

The importance of the paper by BKZ is that they argue for an entirely different approach. Instead of modeling the data generating process (DGP), they assume it is complex and interactive, and that prior efforts to model the origin of conflicts using game theory have not amounted to much (at least not anything testable). They conclude that the only reasonable standard for evaluating a statistical model is out-of-sample performance, without regard to the assumptions or specification of the statistical model. Not surprisingly, they adopt a non-parametric approach and use a neural network to generate an empirical model of conflict without regard to any underlying theory. Their article thus challenges the current methodological orientation of the discipline, insofar as they eschew the ideal of mapping a strategic game to an empirical specification.

I was confident that BKZ were wrong on several particulars, most notably whether their model actually outperformed the standard logit model used by many scholars in quantitative international relations. Along with Christopher Gelpi and Jeffrey Grynaviski, I wrote a reply addressing this problem. Additionally, we presented a general framework for comparing models when the goal is maximizing out-of-sample performance.[1] The larger epistemological questions raised by BKZ remained, however, and their challenge cast into doubt the proper relationship between deductive and empirical models. This dispute and how it relates to the broader themes of this book are dealt with starting in the next chapter.

The final source of inspiration that led to this book concerned an ideal of the proper approach to mathematical modeling in the social sciences. This ideal was first advanced in a set of workshops dubbed "Empirical Implications of Theoretical Models" (hereafter, EITM) funded by the National Science Foundation (NSF) in 2002. After these initial meetings, EITM evolved into a joint effort of Harvard, Michigan, Duke, and Berkeley to train advanced graduate students during the summer. Unlike other methods workshops that focus on particular skills (e.g., the Interuniversity Consortium for Political and Social Research's summer courses), EITM has the larger, epistemological goal of helping young researchers to bridge the divide between

[1] Our article, plus a response from BKZ, is in the May 2004 issue of the *American Political Science Review*.

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521853621 - Computational and Mathematical Modeling in the Social Sciences
Scoot de Marchi
Excerpt
More information

deductive and empirical methodology. The goals of EITM were summarized in a 2002 report presented at the NSF:

> 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.

Although out-of-sample forecasting is specifically emphasized in the above passage, it is obvious that the EITM founders have in mind something quite different than the nonparametric work of BKZ. Their goal is to rework the discipline so that the chasm between formal modelers and empirical researchers is bridged, with the hopes that this synthesis will lead to better models that have clearly testable empirical hypotheses.

By and large, I was very sympathetic to the goals of EITM, and was lucky enough to be invited to participate as a faculty member in the 2003 session at Michigan. My job seemed easy: take two days and present a framework for accomplishing EITM-style research. In my mind, this meant making an argument for how one might bridge the gap between models (usually deductive) and empirical tests; currently, the clearest statement of the difficulties inherent in this problem is

CAMBRIDGE

Cambridge University Press
0521853621 - Computational and Mathematical Modeling in the Social Sciences
Scoot de Marchi
Excerpt
More information

found in two articles by Signorino (1999) and Ramsay and Signorino (2003). After a bit of reflection, the issues involved were more difficult than I at first realized. Many of the arguments presented in this text are a direct result of the questions I faced in formulating my talks for EITM. Chapter 3 lays the groundwork for this investigation, and Chapters 4 and 5 provide a set of tentative answers to how one might implement the EITM statement on methodology.

WHAT THIS BOOK IS NOT

Before proceeding, it is important to say what this book is not. This book, despite appearances in some places, is not a critique of game theory (or formal theory more broadly). Although I am critical of some current practices, it should be obvious that I firmly believe in the aspirations of those who wish to make political science an actual science, complete with predictions and policy advice about events in the real world. My main concern is that game theory has become confused with definitions of human rationality. In this text, I will argue that game theory is a mathematical tool, not a proxy for human rationality where if one departs from game theoretic models one automatically sacrifices any notion of rational agents. As a tool, it is one way to "solve" problems and is better suited to some classes of problems than others. Most of the examples I focus upon concern classes of problems that for a number of reasons are ill-suited for a game theoretic approach, and I propose a set of methods "rational" agents might employ to deal with these complications. The reason for providing tools that expand the class of problems one can deal with analytically is in my mind simple: better models, with more verisimilitude, allow an easier transition to empirical tests. This is the primary goal advanced by EITM.

Game theory also has been confused with pure mathematics, insofar as many practitioners feel no need to connect their models to empirical tests. Much that masquerades under the classification of "theory building" is not worth the appendices, and one should question the usefulness of models that rely upon limiting assumptions to produce whatever narrow result is desired by the researcher.[2] Following Granger (1999),

[2] Arrow's impossibility theorem, in contrast, depends upon assumptions that are of substantive interest and produces a result that is extraordinarily broad.

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521853621 - Computational and Mathematical Modeling in the Social Sciences
Scoot de Marchi
Excerpt
More information

the viewpoint adopted here is that the connection between theoretical models and their empirical referents needs to be direct enough such that we can be satisfied that the tests we conduct are actually dispositive. Dispositive tests distinguish the actual model (or data generating process) from the universe of possible models. This viewpoint is by no means new; rather, it has been the subject of debate in economics for decades.[3] What is perhaps new will be the particular modeling approach adopted here, which combines traditional game theoretic investigations with computational models. The reason for this union hopefully will become clear in subsequent chapters.

This book is also not a critique of empirical work in political science, though, again, one might be confused given that in places I am critical of existing efforts. Just as models without empirical tests are suspect, so, too, are data-driven statistical investigations that fail to make apparent what model is being tested. Good statistical work allows us to distinguish useful models from the universe of irrelevant models; further, it allows us to investigate the generality of a model and the places where assumptions are carrying too much of the load. I will, however, place rather more emphasis on predictive work than is currently the norm within the social sciences, as much of the statistical research that has been conducted in the social sciences aims solely at comparing the in-sample performance (or "explanatory power") of various models.

In-sample comparisons should be seen as innately suspect, as one can easily overfit a statistical model and claim "success" for a theory. More time will thus be spent in this text addressing the curse of dimensionality that has to this point been largely ignored by social scientists.

A Simple Example: Applause, Applause

As is fitting for a book on modeling, let us begin with a simple question. Hopefully, this will introduce most of my essential arguments before we wade into the deep end of the book. The history of this example is

[3] See, for example, the October 1993 *Special Issue Anniversary of the American Journal of Agricultural Economics*. Castle (1993) and Leontief (1993) are particularly useful in this issue, insofar as they outline a set of requirements that would help connect deductive models with empirical tests.

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521853621 - Computational and Mathematical Modeling in the Social Sciences
Scoot de Marchi
Excerpt
More information

a rich one, given that it was used for many years by John Miller and Scott Page at their Computational Economics Workshop at the Santa Fe Institute.[4]

Imagine you are asked to explain or predict the occurrence of standing ovations. You have a performance of some type, where each member of the audience receives a signal from the performance about how good it is (based upon their own internal preferences). Each audience member can then choose to do nothing, applaud, or stand and applaud. They also can sit down again at any point should they decide to stand initially. This is a highly stylized problem but has relevance for social scientists. We often want to understand who stands, or votes, or participates in a riot, and how individual characteristics and social dynamics lead to this behavior.

There are different approaches one might take to this problem, and in social science one can roughly describe the three methodological traditions that could be utilized: empirical, deductive (i.e., game theoretic), and computational. Let us investigate what sorts of answers these traditions, in isolation, might provide to the standing ovation problem.

An empirical researcher would likely start out with questions concerning what measures would be collected for both the dependent and independent variables, and not all of the forms of these measures would be obvious. For example, the dependent variable might be coded as a binary variable measuring whether or not the ovation occurred. If this encoding is adopted, what would the right threshold be for distinguishing an ovation? Would 90% have to stand? More? Less? The choice of scale for the dependent variable is also not obvious; one could change both the temporal and spatial characteristics of the dependent variable. For example, one encoding would measure the length in time of the ovation, but any such measure of time would retain the problem of choosing an appropriate threshold. Alternately, one could measure the likelihood that any given audience member participates in the ovation, thus changing the unit of analysis spatially from the entire audience to each individual member.

[4] Past answers to this problem are archived at http://zia.hss.cmu.edu/econ/homework95.html. For the most recent investigation of this problem, see Miller and Page (forthcoming).

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521853621 - Computational and Mathematical Modeling in the Social Sciences
Scoot de Marchi
Excerpt
More information

Another more insidious problem would involve the nonindependence of observations.[5] Cleary, if subsets of the data set involved repeat performances by the same artist, "buzz" might result in a lack of independent, identically distributed (IID) observations. This problem also would complicate the measure of independent variables. Measures of performance quality and the like could easily be contaminated by interactions either between guests of the same performance (e.g., social pressure) or for members that attend multiple performances across observations in the data set. And members of different audiences are obviously not drawn from identical distributions, as people sometimes choose which performances they attend.

Problems aside, what sorts of questions would the empirical researcher answer? Likely, it would involve establishing relationships between such concepts as "performance quality" (as perceived by the audience), the type of performance, the number of audience members, and so on, and the likelihood or length of an ovation.

A deductive (or formal) modeler would come at this problem from a different angle, where the most important decision would involve specifying the benefits and costs that are present for members of the audience when they decide to ovate or not. Clearly, you do not want to be the only fool in the audience standing and clapping madly; people would stare. Just as clear, you do not want to be the grinch, sitting alone in a sea of excited fans. At some level, though "quality" matters, you only want to reward "good" performances with an ovation, given the effort involved in standing and clapping.

The structure of the game would also involve a set of important considerations on the part of the deductive modeler. How many periods would be included in the game where agents could update their information and make choices? If an ovation occurred, how would people get back to their seats? The same sorts of utility considerations discussed in the preceding paragraph would apply with equal force to agents making choices to sit back down again.

Given these modeling choices, and the input of a few "state of nature variables" such as the quality of the performance, the deductive

[5] One also might point out that the observations are not independent spatially – that is, whether or not one member of the audience stands (or later, sits) is likely correlated with the actions of other audience members.

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521853621 - Computational and Mathematical Modeling in the Social Sciences
Scoot de Marchi
Excerpt
More information

modeler might well reach a good understanding of the individual decisions that work together to produce an ovation. A model might also help worried performance-goers in reaching decisions about whether or not to stand for an ovation in future performances. Ever present, however, would be the worry that the limiting assumptions relied upon to formulate a sufficiently simple model might cut against the usefulness of any insight gained.

The final tradition that might generate a solution for this problem is less well known in the social sciences. A computational (or dynamic systems) researcher, in contrast to the two preceding approaches, would specify a set of rules that governed the behavior of individual audience members, along with a set of contextual variables that described such features as the seating arrangement, the shape of the performance hall, relationships between audience members, and so on. What would these rules look like? On one level, the rules would be functional expressions that would be similar to the utility functions used by a game theorist, though these functions might well be allowed to vary both in time and by the individual type of audience member. On another level, these rules could add substantial verisimilitude to the computational model by incorporating features of the problem that would be difficult to model in a deductive framework (e.g., learning models based upon research in cognitive psychology). One such rule might involve adding vision to the model – given the shape of the performance hall, not all audience members can physically see all other audience members. Any utility function that involved peer pressure should be more sensitive to people within an agent's field of vision than agents outside this field.[6]

Unlike a game theoretic model, it is unlikely that a computational model would produce a set of deductive results. What is far more likely is that the researcher, confronted with the large parameter space generated by the rules used in formulating the computational model, would have to rely upon statistical investigations to understand

6 The outcome of such a rule is that not all audience members are created equal – that is, audience members in the middle rows nearest the stage would have a disproportionate share of influence. One also might consider the type of individual audience members. For example, if a group of Catholics got together to watch a play, it might matter if the Pope were sitting in the audience. I would hazard that if the Pope ovates, so, too, would everyone else.

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521853621 - Computational and Mathematical Modeling in the Social Sciences
Scoot de Marchi
Excerpt
More information

any "results" of the computational model, much as in the empirical tradition. Statistical relationships between parameters, rules of interest, and the likelihood of an ovation would then be presented, albeit substituting artificial data for real data.

This is a brief sketch of an interesting problem, but it raises questions of importance to all modelers. To begin with, are the approaches complementary or distinct? On the face of it, our three stereotypical methodologists would not have much to say to each other. The empirical researcher is establishing correlations between different measures and the likelihood of ovation; the game theorist provides advice on how rational audience members should select strategies; and the computational modeler incorporates aspects of both of the forgoing approaches to produce a dynamic model that recreates a standing ovation.

All of these models ostensibly explain the same phenomenon, but can one compare or integrate the results? Or, are these simply different answers to different questions? I will argue in the succeeding chapters that it is undesirable to let each type of modeler work in a vacuum; models need to produce results that are directly comparable to competing explanations. Even within each methodological approach, models are often not unique. Different modelers will produce different answers, and the job of social science should be to sort among them by insisting on out-of-sample tests of some kind. If, for example, we are confronted with several different game theoretic models, all explaining standing ovations, how do we decide which one is closest to being right? Unless one of the game theorists makes a deductive mistake, the models will differ because the assumptions differ. Arguing about assumptions is a little like arguing about whether Wolverine is tougher than the Hulk; ultimately, it comes down to taste. This book will argue that a different, integrated approach is required to make sense of these questions.

STRIFE BETWEEN METHODOLOGICAL CAMPS

Currently, there is a sense of mutual distrust between different methodological camps. Let us start with the more forceful critiques of the empirical tradition. As part of the EITM meetings, Christopher Achen argued that one must be suspicious of empirical modeling in the social sciences. Because many models are quite complex, researchers have

© Cambridge University Press www.cambridge.org