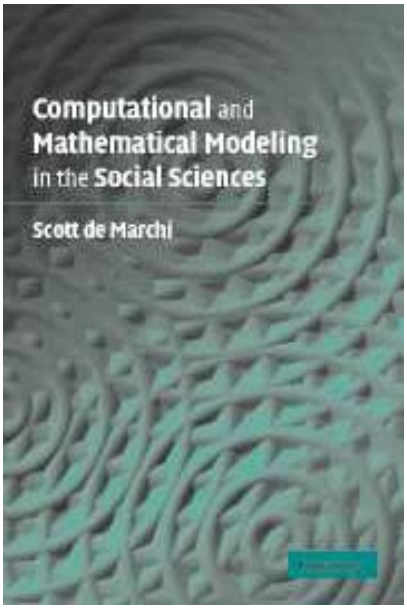
Computational and Mathematical Modeling in the Social Sciences

Marchi, Scott de
Cambridge University Press: Cambridge, 2005
ISBN 0521853621 (hb), 0521619130 (pb)

Order this book

Reviewed by **Flaminio Squazzoni**
Department of Social Sciences, University of Brescia, Italy



The road to science is paved with sacrifices. Sacrifices to learn and pursue rigor, under the strict constraints of any formalism. Sacrifices to collect empirical data and to search for empirical foundations of models. Sacrifices to fine-tune theories and empirical findings. Sacrifices to avoid the sirens' song of social science philosophers, theorists, and pamphletists, who never did analytical science and empirical research.

This book can be seen as a kind of homage to real social scientists, as well as good bread for daily use. In fact, the author doesn't limit himself to argue in favour of an analytical approach to social sciences, but suggests a framework for a unified model-based approach to social sciences, trying to combine game theory, statistics and computational methods. The author rightly argues that, even if most of scientists would see these as different methods or even theories, formal models, statistics and computation should be viewed just as different tools of the same scientific enterprise. The author's purpose is both to defend the social sciences from the anti-positivistic and post-modernistic drifts, and to improve the capacity of science to tackle complex problems and challenges, increasing, at the same time, theoretical transparency and empirical foundations of models. In doing so, he strengthens the foundations of a computational approach to social science. It is good to see that he tried to do it with a clear-cut and well written book, with examples, intuitions, and personal anecdotes.

The book consists of six chapters, with a prelude, where the author summarises his viewpoint and the book's purpose. Right from the prelude, the author clarifies both successes and pitfalls of the mathematical and standard approach to social science. He argues the need for a revision of the mathematical approach, such as game-theory and statistical models, since "existing methods are brittle when confronted with complex problems, and there is a genuine lack of correspondence between deductive models, on the one hand, and empirical tests of these models, on the other" (p. *xvi*). Critical examples cover both technical and methodological problems, such as the "curse of dimensionality", or theoretical problems, such as the consequence of the simplified assumption of the individual preferences on which each standard model is built. But, his position should be seen as a constructive critique of existing practices, with the aim of improving the way science works. The main challenge can be summarised as follows: "how to build more complex models of behaviour without sacrificing the ability to subject the results to exacting scrutiny?", at the edge of theoretical abstraction and empirical contents (*xx*). The answer relies upon the combination between formal models, computational tests and empirical foundations.

The first chapter (*Not All Fun and Games. Challenges in Mathematical Modeling*) summarises the overall inspiration for the book. The author clarifies that the book is not intended as a critique of game theory or formal theory and modelling in social sciences, as well as a critique of empirical analyses, but an attempt of

combining the methods. After a close confrontation of the respective trade-offs of the three methods (formal, empirical, computational), he concludes that "a structured combination of the methodological approaches is so far superior to any approach taken separately" (18). The rest of the book can be seen as an attempt to defend this argument.

The author's suggestion is to fill the gap between empirical and formal models, the first ones interested in detail and specificity, the second ones in abstraction and generality, using computational models to explore the implication of a purely deductive theory and to allow the scientists building models with more verisimilitude. This means the creation of a "feature space", or an "assumption space", between analytical models and empirical tests, which should be carefully explored via computational models, applying the statistical approach as a kind of methodological analogy. Computational techniques add a fundamental advantage to this process: exploring the space of possible assumptions of analytical models with a reduction of time and resource constraints.

The second chapter (*Looking for Car Keys Without Any Street Lights*) focuses on methodological challenges in building empirical models. The author emphasises the over and underfitting of the models, the problem of the curse of dimensionality, the exploration constraints in the size of the parameter space, and the impossibility of testing theoretical models with statistical methods. At the same time, he also emphasises the risk of overfitting game theory, formal and deductive models. Actually, just as empirical models can overfit data, so deductive models can overfit desired results. This is particularly evident in the case of game theory models. The conclusion is that social scientists should work to a "theory of *equivalence classes* for formal or game theoretical models" to avoid the above mentioned methodological risk (73).

At a close outlook, it is worth noting that the chapter is founded on an extensive literature on methodology of social science, and on a selection of examples, such as the prediction models of conflicts between nations and the currency game, driven by the author's background, which is political science and, in particular, computational political economy and public policy. To be honest, one would argue that other examples would be more effective, that the brevity of the chapter does not guarantee enough space to enter in details as one would expect, or that examples are simply suggested as crutches to set the methodological excursus up, without entering in real and concrete applications. This can be true. But, the good point is that the author's argument is clearly defined and is a logical prosecution of the reasoning presented in the first chapter.

The third chapter (*From Curses to Complexity. The Justification for Computational Modeling*) is where computational method enters the picture and the author develops the justification of his unified approach. The example used is international conflict, in particular the conflict initiation, with a careful (positive and constructive) critique of game theory modelling and the proposal of how to complicate the picture to improve the theoretical value of formal theories, via computational modelling. The constructive points for going beyond the limit of standard modelling are some methodological steps to "allow researchers to explicitly model more complex games and build upon prior effort using feature spaces and domain-specific encodings, avoid brittle encodings that limit the generalizability of results (i.e. develop an equivalence class for your model), introduce a methodology to model component games and their associated (idiosyncratic) utility functions" (96). Game theory can be used to develop intuitions about the scientific problem, computational modelling and optimisation theory should be employed to identify and "traverse the space of possible components and idiosyncratic utility functions", and, last step, the results of the model should be tested for stability against "similar games or against human play" (111). This is to provide the equivalence classes of games, which the author conceive as a link between deductive and computational models, on the one hand, and empirical referents, on the other.

The fourth chapter (*Why Everything Should Look Like a Nail. Deriving Parsimonious Encodings for Complex Games*) aims at providing some preliminary foundations for scholars interested in computational modelling and techniques. Questions addressed are as follows: how to decide about the programming language, how to develop the environment, and how to build a computational model, from a simple to a complex structure. To the JASSS readers, this could be viewed as a simplistic picture, summarised in a few pages, and that's it. But, the point is that the chapter is addressed to scholars who are not experts in programming languages, maybe scholars more inclined towards statistical tools and game theory. This is the reason why the chapter, as well as the entire book, can promote the computational approach in the standard social science.

The fifth chapter (*KKW Redux. Deriving and Testing Logical Implications*) deepens the challenge of

choosing models from the class of locally consistent models using out-of-sample comparisons. KKW stands for a methodological contribution by Gary King, Robert Keohane and Sidney Verba, which the author follows to emphasise the importance of deriving the logical implications of models. Following in these footsteps, the author clarifies that there are a few examples in the social sciences where logical implications are used as an alternative means to statistical modelling to testing a model. In spite of this, the author uses examples from the study of voter preferences, and intuitions from Kauffman's NK models, and other complexity models, to model preferences and derive a measure for survey data, even in the form of rough approximation able to capture the essence of full measures. From the example taken from voter preferences, through the application of Kauffman's NK approach, the author is able to derive the logical implication that the variance in preferred policy platforms in each subgroup is a proxy for actual complexity in the electorate, with the consequence of suggesting a simple and parsimonious measure of the electorate's complexity. This is just an example of how to test the main implications of a model by deriving logical implications that allow matching theories and data, but an enough simple and convincing example.

The closing chapter summarises some important conclusion. The author goes back to his starting points. He reminds that the main methodological problem to be addressed is the curse of dimensionality: "when one resorts to limiting assumptions in game theoretic work, or distributional choices and a functional form in empirical work, one is implicitly making a statement about the parameter space" (176). To tackle this methodological challenge, one should, first, derive a "feature space", enough appropriate to the research question, so that details and empirical data can be incorporated in theoretical models. Afterwards, one should resort to out-of-sample testing (data or logical implications), given the fact that evaluating models through in-sample comparisons is impossible. Last step, one should develop more complex models, to further incorporate real-world processes in the model, via computational models able to link deduction and empirical foundations.

In conclusion, the book makes several positive points, also for JASSS readers. For people outside computational social science, the advantage is that the book presents the computational modelling as a fundamental part of the scientific process, and allows standard social scientists becoming acquainted with it. Despite some critical point, such as the brevity of the book that often implies lack of details, the need for an in-depth examination of some relevant problem, such as the quest of how to explore the space of possible models, given time, resource, and theoretical constraints, the exclusion of other empirical methods beyond statistics, and so on, for JASSS readers, I see the following positive contribution: frequently, social simulation practitioners tend to overrate the difference between computational, empirical and mathematical models, as though these models were different scientific enterprises, while, differently, this book allows combining methods rather than discriminating, so that differences in methods and tools are seen as a positive property of the same unified scientific enterprise. That's not a trifle.

Return to Contents of this issue