PARAMETERIZE!: Notes on Mathematical Modeling for Sociology

HARRISON C. WHITE
Columbia University

Radius is parameter to circles. Mathematical modeling contributes best to sociological theory through parameterizing. Parameters form a seam linking quantitative and qualitative panels of theory. All social construction comes through feedback, and modeling its outcomes requires use of parameters.

Mathematical modeling is one of many routes for developing sociological theory. Like each other route, it is exploited best in combination with others from among historical, survey, exegetical, linguistic, whatever—so as together to uncover and make sense of phenomena. I argue that the distinctive contributions to theory from math modeling, as contrasted, say, with those from statistical fittings of quantitative data, derive from its search for and attention to *parameters*. The radius is parameter for circles. I shall make two main points about these searches.

My first point is that parameterization offers ways to couple qualitative with quantitative analysis. Using tense or case in analysis of text, for example, can be seen as parameterizing, and so too for grammar on larger scale. Yet, equally, a parameter "characterizes a member of a system (as a family of curves)" to quote from Webster's, and so is all bound up with calculating motions and profiles. Note that in both perspectives the parameter is on a higher level, and mathematics does indeed offer exceptional help with sorting out into levels as a control over understanding the story, which amounts to supplying theory.

By themselves, quantities just aid assessments, but when parameters marshal them into an array of numbers, then they can trigger alternative thinking. Mathematics can even permit bypassing the numbers. For example, the mathematics of symmetry (called abstract group theory) supplies the parameter 5 for a pentagon without much bother to verify invariance under those rotations using physical trials.

The task of theory is to probe beneath appearances. It took humans millennia first to build a familiar (and of course partly accurate) view of the sky as a Milky Way, and then it took centuries for theory to probe to a causal level, expressed parametrically. A Milky Way configuration also is drummed into us sociology theorists by the folk wisdom we necessarily absorb. It is called variously Society, State, Economy. Mathematical modeling surely can help us also to penetrate beneath such common sense. As with astronomy, it must be a stripping operation, in order to get an alternate picture clear enough to be falsifiable and so feed back into further probes. Society, State, Economy will prove to be as amazingly different from what they shadow as Orion and his folk-view cousins are different from black holes, X-ray clouds, and other exotica known to be in the sky.

Yet, social constructions will also prove amazing in a different way, since they are inter-active and thus must induce interpretability along with agency. The variety of historical paths by which they emerge must be observed with care and will require some quite new forms of parameterization. Indeed, an example of parameterizing feedback will be used below to support my second main point. There also must be periods of involution in this model building, self-absorption lasting long enough to gear it up, intermeshed with a coterie of practitioners growing deft in application.

Sociological Theory 18:3 November 2000

So much for rhetorical introduction. My notes on the actual talk I gave at the ASA 1999 panel organized by Tom and chaired by Willie are a bit scrappy. By instruction of the Editor, this written version is to be illustrated with examples from relevant research work of my own. So I will interweave such examples in and then return at the end to the most general remarks I made. (Exact cites to these earlier publications of mine can be found in my 1992 *Identity and Control* from Princeton University Press.)

Let me start out with a sketch of my own research background that can clarify how parameters look in the wild and how to tame them. My first explicit thoughts along these lines date back half a century to my first two years at M.I.T., when I was exposed to (very good) instruction in literature and history while continuing to build up the math and natural science of my high school days. My thoughts were often of chains of effects spreading among persons around the world, of how to penetrate the mysteries of cause and effect. At some point, perhaps when studying with Karl Deutsch at M.I.T. while he was writing *Nationalism and Social Communication*, the answers came to seem likely to emerge as social structures constructed among us, by us.

Deutsch suggested switching to social science. I pondered that and explored some social science while completing my physics Ph.D. The Ford Foundation offered a "switching" fellowship which I took to Princeton. There Marion Levy resonated with my commitment to theory and Wilbert Moore did so as to history, but it was Fred Stephan and Frank Notestein whose practice in research reinforced my commitment to parameterization.

Vagaries of career led me through a job in Operations Research with the army, where I worked on inducing battle orders and unit identification from cross-referencing traffics of all sorts, in signals and field exercises . . . Then a blessed year at the Center for Advanced Study in Stanford let me steep myself in a series of English anthropologists—with E. E. Evans-Pritchard there in the flesh—whose work came closest to my intuitive sense of how social reality came into existence. Higher-order numbers, parameters, often were what distinguished people's ways of organizing themselves.

Finally, launched on my first academic job in the Carnegie-Mellon business school, I could actually carry out my first social study. It was of management structure in a (now defunct) heavy manufacturing firm, Firth-Sterling, British-owned and trying to get into modern times through research and development. Sure enough, math modeling, in the form of multiplex sociometry as combined with balance theoretic modeling of attitudes, enriched interpretations from my codings of linguistic register and departmental infighting. The books of James D. Thompson and of March and Simon helped me to shape these into a coherent theoretical vision, one characterized partly by sociometric parameters.

Later, in the sociology department at University of Chicago, I could take the time, with encouragement from colleague Peter Rossi, to work up a structuralist model of entire social anatomies—of various Australian aboriginal tribes which cohered around very specific systems of classificatory kinship. Mathematical representation theory of abstract groups (cf. the pentagon earlier) offered leverage. Since this nonquantitative mathematics describes and interrelates possible interchanges of entities, it is a natural framing for the exchanges which build kinship.

This mathematics suggested a number of possible homogeneous, yet highly intricate, role systems that I could check against an array of field monographs. One of the best of these was by an American sociologist, W. Lloyd Warner. My first book was a marriage of the two. Later, a brilliant relaxation of the group theory constraints that mandated closure in the system of clans was developed by John Paul Boyd. He made excellent use of the mathematics of semigroup representations. John had been trained in one of the novelty science disciplines that emerge from time to time in the ferment of Ann Arbor (it's a long, long winter!). His work emboldened me (by now at Harvard) to operationalize, jointly with young collaborators, the underlying social logic of sociometric networks in our own current world. My doctoral training in theory of the solid state, also at M.I.T. in physics, had prepared me to apply representation theory of groups and thus also of semigroups.

Earlier at Chicago I had, on the side, also worked at a complementary vision of social pattern as emergent from chance rather than from symmetry. I proposed a combinatoric calculus for social mobility tables.This is a case where modeling (at least mine) was supplanted by a powerful framework for statistical fittings, from Leo Goodman—although Joel Levine is now proposing a yet more powerful mathematical model that parameterizes. The combinatorial model had been stimulated by my training in statistical mechanics; so also were my queuing studies, which suggested clues for understanding of many social processes.

Still, my production was coming to seem a bit of a Chinese dinner, tasty but not properly filling. *There wasn't a satisfactory parameter in any of that work.* I did not count integers from group theory such as the order of permutation operators in kinship cycles. This was because my intuition said that to be most helpful a parameter should be *tunable*.

I had been haunted from M.I.T. days by a classic little monograph from the 1920s authored by physicist Percy Bridgman (who won a Nobel for studies of matter under extremely high pressures). In stunningly clear prose, with but a few simple and elegant equations, Bridgman hammered at this *Dimensional Analysis*. From the beginning I disdained statistical "fitting," which produced numbers which had stability neither in size nor in meaning—numbers which were, exactly, not parameters. Bridgman indeed upped the ante; Bridgman sought superparameters, numbers which held across vast stretches. Some of these identify scaling laws. I learned that George Kingsley Zipf in 1939 published an extensive, if simple-minded, search for similar superparameterization of social phenomena, in his (and earlier in Pareto's, and later in Simon's) distributions by size and/or ranking (of cities, words, money, just about anything).

My impetus, more prosaic than that, was toward just ordinary parameters. In the course of searching, I did also finally manage to scratch my original itch, from sophomore year, about causal chains. I developed a vacancy chain approach to job mobility. The archival work required to dredge up data proved absorbing, but I was especially energized by the thought there would be a path to tunable parameters from complete and accurate data. But, alas, my theoretical hunch about tenure network involution has not, at least not yet, showed up from such data as I coded; I was left only a half-step advanced, in the form of Markov Chain probabilities, as over against statistical fitting constants, despite the better understanding of chains of causation that was achieved.

So, after developing these vacancy chain models, as well as the sociometric blockmodels from semigroup theory, I turned to modeling operations of production markets where parameters indeed could be made absolutely central. Parameterizing enabled me to develop a map for locating sorts of markets along with expected outcomes (profits, size, all that). My market work was interspersed with attempts to get at control processes, however, and did not produce a full-length manuscript account until this past January. It is awash in use and presentations of parameters, guiding Figures and Tables alike.

Oddly enough, economic theorists of markets, who talked as if they were doing physics-like analysis, disdained parameters. They turned out to resemble in my eyes a Melanesian cult. Led by Paul Samuelson, this cult had happily accepted mysterious cargo dropped from field theory "airplanes" in physics without ever understanding that parameterization was indispensable for tying to realities. These economic theorists just wrote EQUATIONS, perfectly general and remarkably useless equations. Like Melanesian chiefs they—aside from odd-men-out like Wassily Leontief—were not worried by the lack of operational results. They provided the religious incantations (equations without parameters) while a lower caste (Industrial Organization practitioners and the like) laboriously, and behind the scenes, did drudge work in description. The latter were unguided by theory and did not seem exhilarated by the screed of Holy Equations spread across their economics firmament.

In my new manuscript, the amount of production from each one of the half-dozen or so firms in an industrial market depends on the amounts from the others. So there is feed-back between their sizes and the overall aggregate size of the market, mediated by the firms' ranking by quality in eyes of the buyers downstream as well in their own eyes. Yet, also, that overall size, say for the jet engine market, is affected by the corresponding sizes for parallel markets, such as for turbo-prop engines. Implications of this latter feedback can be captured in a cross-sectional model. A key role is played by a parameter for substitutability. This is a single number whose deviation from unity reflects the feedback between competing tendencies to growth, exponentiated. I call it gamma. For the engine case gamma might rise to close to unity (independence) as jet engines became more predominant. In a cooking-war sector of industry, glass and ceramic pot markets might have gamma small. And of course gamma might be larger than unity among various software markets in this computer era.

Even with this parameterizing, the modeling of markets still seems incomplete. Further depth requires seeing how these interlocking choices get to be framed in perceptual terms that are grounded in patterns of discourse and thus culture which themselves resulted from accumulated history of these social construction of markets.

Parameterization surely is needed, as claimed in my introduction, and especially with analysis of language and other formats of interpretation. To parameterize is both path to and sign of analysis that is going beneath surface appearances. This vision must, however, be generalized in order to interface its analytic mode with pragmatics of action. Consider an analogy. Without metallurgy, the analogy goes, never would the instruments for penetrating the Milky Way have come into existence, and yet it is hard to think of a field less helped during its history by theory and parameterizing than metallurgy. Sometimes applied work uncovers on/off distinctions but much more commonly the issue is exact degree, and there some version of statistical fitting does come into its own, as in the seminal work of Leo Goodman, mentioned earlier, and of econometrician Zvi Griliches, who taught us about diffusion of hybrid corn as well as of features of Detroit autos.

Mathematics as we know it is largely offshoot, however indirect or concealed, from natural science theory; so considerable extension and adaptation may be needed, even as to pure theory apart from statistical fitting. This need is especially vivid with respect to space. Space is the central construct of all the natural science, though this is not visible to most practicing scientists.

Two decades ago the great physicist Kenneth Wilson was finally penetrating the mysteries of physical dimensionality at a level yet more fundamental than Bridgman's. Wilson integrated combinatorial aspects with field theory aspects to uncover, exactly, superparameters. He derived dimensionality as a decimal number, thereby providing deep theory for the scaling parameters that had been being established on *ad hoc* bases across any number of physico-chemical contexts, and, as noted earlier, for some social contexts. If I were younger, I'd try to figure out how to unpack this very difficult analysis so as to make it available for socio-cultural processes.

Social networks are one, still primitive initiative toward conceptualizing the analogue(s) of space for socio-cultural process. Central to this network effort is the concept of structural equivalence, expounded in Ronald Burt's works, which is surely as distinctive for the socio-cultural as dimensionality is for the physico-bio-chemical spaces. Other such constructs, ones more combinatory and stochastic, also are central and also are handled differently from in other sciences.

Time is the real puzzle. This is true even for natural science, but put that to one side. Humans are not responsible for space in anything like the extent to which time is the entail of humanity. Story, interpretation vehicle, is our instantiation of time. Language, in the pragmatic sense of discourse process, has emerged from and with story/time. Multiplexity is central to time thus playing out in interpretation of socio-cultural process. Cross-cutting, not cumulation, is the central moment. Contingency characterizes social time, not as arena for resolution into outcome but instead for the emergence of theme that will re-emerge again and again, both in interactive fact and by interpretive construction. The usual modeling formulae for change, in differential and partial differential equations, are inadequate for social science. This is already appreciated in qualitative theory, as in books by Niklas Luhmann and the systems theories on which he keyed, but mathematical modeling lags. Very simple tools can permit progress if brought to bear on tangible institutional systems about which people care enough to generate rich and varied accounts, "data." These tools include Venn diagrams along with their generalizations by theory lattices, which have been applied to socio-political mobilization by Ann Mische.

The encouraging fact is that progress is being made exactly through mutual informing between modeling and theory. The so-called micro-macro gap is an optical illusion of trying to visualize socio-cultural process through natural science framings. "Middle Range Order" is what is supposed to be missing. But this vision is based on physical, metric space and time. There is no such "middle range order" for socio-cultural process, which instead is concerned with interpenetrations that are at once micro and macro. Results supporting this view are beginning to appear, for example, in anthropological linguistics, and perhaps this is what Anthony Giddens is getting at in the books he dashes off. And this view is being reinforced by results of a new genus of simulations, such as by Kathleen Carley and by Duncan Watts. They are made possible by new sophistication and by exceptional increases in computing power. These simulation studies are woven in with and around explicit mathematical modeling and parameterizing rather than being exercises in brute force.

REFERENCES

Boyd, John P. 1991. *Social Semigroups: A Unified Theory of Scaling and Blockmodeling as Applied to Social Networks*. Fairfax, VA: George Mason University Press.

Burt, Ronald S. 1992. *Structural Holes*. Cambridge, MA: Harvard University Press.

Carley, Kathleen. 1992. "Extracting, Representing, and Analyzing Mental Models." *Social Forces* 70:601–36.

Mische, Ann. 1998. *Projecting Democracy: Contexts and Dynamics of Youth Activism in the Brazilian Impeachment Movement*. New York: Doctoral Dissertation, New School for Social Research.

Muir, James. 1972. *A Modern Approach to English Grammar: An Introduction to Systemic Grammar*. London: Batsford.

Silverstein, Michael. 1993. "Metapragmatic Discourse and Metapragmatic Function." Chapter 2 in John A. Lucy, ed., *Reflexive Language: Reported speech and metapragmatics*. New York: Cambridge University Press.

Watts, Duncan. 1999. *Small Worlds*. Princeton, NJ: Princeton University Press.