

Simplicity, generality and truth in social modelling

Part 1: Epistemological issues

Scott Moss

Centre for Policy Modelling, Manchester Metropolitan University Business School
`scott.moss@scott.moss.name`

Abstract. The paper addresses the standard issues of simplicity, generality and truth as they apply to social modelling. It is argued that evidence driven, agent based social simulation turns on their head constraints on model or theory simplicity and leads naturally to empirical validation (hence assessment of truth). Generalisation follows by abstracting elements of empirical models. A research programme of empirical model and model abstraction could form the basis for and empirically based, general social theory. A demonstration of the elements of this programme is reported in Part 2.

1 Introduction

As in any human endeavour, social modellers tend to conform to a range of social norms. These include constraining model design by appeal to some social theory, designing models to have the smallest possible number of degrees of freedom and producing results that can be interpreted as useful indications about the nature of social processes and/or useful guides to public policy. It is by no means unusual for social theories and analytical techniques to be closely related. Properties and implications of n -person games and general economic equilibrium models, for example, are derived by appeals to fixed point theorems; macroeconomic models are closely bound up with econometrics. Moreover, it is not uncommon for assumptions to be made “for the sake of simplicity” which means that the problem being addressed is modified to render it amenable to analysis with the mathematical or statistical technique associated with the chosen theory. The point about degrees of freedom is that the fewer the degrees of freedom — the smaller the number of “free” variables — the more definite are the social and policy implications of the model.

An alternative approach is evidence-driven agent based social simulation, sometimes called *companion modelling*. The norm in this tradition is to engage with stakeholders and independent domain experts to identify behaviours of the relevant actors and their social context. Models designed and implemented within the constraints of this evidence are used to generate formal scenarios that are presented to the relevant stakeholders and used to inform the development of narrative scenarios for the exploration of alternative social policies or strategies.

Such evidence-based modelling has been undertaken at a range of scales such as within river catchments in Senegal [Barreteau et al., 2001] and Thailand Becu et al. [2003], within a large enterprise [Moss, 1998], across southern England (including both rural areas and London) Downing et al. [2000].¹

Evidence-based social simulation modelling with stakeholder participation is a recent development without established social norms. Different practitioners use different programming languages and development environments, different representations of individual cognition and different modelling paradigms. There is no explicit guidance on such issues as the value of model simplicity. Generality has not been much of an issue precisely because the models are developed to describe and simulate specific, grounded social contexts and behaviour. Nor are these models used to forecast future outcomes but rather they are used to explore and discuss possible futures contingent on various policy or strategic actions.

The purpose of this paper is to explore what modelling norms would be functional and how these are different from the observed norms of conventional social modelling.

2 The issues

2.1 Simplicity

Appeals to authority are a substitute for argument, reasoning and evidence. One such appeal, frequently offered by social (especially economic) modellers, is to Occam's Razor – the aphorism to the effect that a simpler model is a better model. There is, of course, a constraint on simplicity. Either, as Einstein is said to have had it, a theory should be as simple as possible but not simpler or, as Occam (William of Ockham, the 14th century who does not appear to have written the phrase) is claimed to have written, entities should not be multiplied beyond necessity, there is a notion of a limit on how simple a theory should be.

There is no doubt that, in the physical sciences, the principle of Occam's razor has been enormously productive and was a key aesthetic consideration in Einstein's development of relativity theory, for example. Matters could hardly be more different in the social sciences.

There is here an important issue of epistemology. The objective of physical scientists to predict physical outcomes of specified events has been shown to be appropriate because it has been broadly successful. The failure of the social scientists systematically to predict social outcomes of specified events has never led to a consideration of why that has been the case but rather to the elaboration of failed theory – frequently in violation of the principle of Occam's razor. [Edmonds, 2007] In general, social scientists seek to exploit existing theory rather than explore for a better theory and, indeed, a better epistemology.

Now a tendency towards exploitation of what is known rather than an exploration of what is not already well known is a common feature of intellectual advance and social development. It certainly characterises the history of

¹ For a review of the techniques employed by the pioneering researchers in the field, see Becu et al. [2003].

technological [Rosenberg, 1976] and organisational [Chandler, 1977] change and business strategy [Penrose, 1959]. However, as exploitation becomes less beneficial and increasingly expensive, progress takes the form of exploration – trying new ways to achieve goals or perhaps new ways of using prevailing knowledge, experience and physical or technological resources. It would seem hard to justify the proposition that, as a result of continued exploitation, social theory has been getting either simpler or better at informing social policy or predicting the course or outcomes of social processes.

So far, I have taken the meaning of the words “simple” and “simpler” in their vernacular senses of being uncomplicated and less complicated, respectively. Many social modellers have what amounts to another meaning of simplicity as a justification for relying on assumptions that distort the reality of the social phenomena they seek to capture with their models. A search of almost any issue of any leading economics journal will turn up at least one assumption made “for the sake of simplicity”. Typically, this means that a straightforward formalisation of observed social relationships would make the application of some economic theory or conventional modelling technique either infeasible or intractable. Assumptions of linearity and normality are common “simplifications”. However, other examples are not hard to find.

To distort the description of a problem or assume it has unobservable properties when those assumptions do not lead to empirically validated predictions or validated specifications of social processes should reasonably render models or theories based on such assumptions inapplicable to policy analysis or social explanation. Such “simplifying” assumptions are such patently bad science that this form of simplicity and simplification will not be considered further here.

2.2 Generality

We shall say that a model G (for general) is more general than a model S (for special) if the set of conditions in which model G is applicable is a superset of the conditions of applicability of model S. A condition of application is positive if it is the specification of circumstances in which a model is applicable or negative if it specifies circumstances in which a model is not applicable. Consequently, a model G will be more general than a model S if the negative conditions of application of the former are a subset of the negative conditions of application of the latter and the positive conditions of application model G are a superset of those of model S and either the negative conditions are a proper subset or the positive conditions a proper superset. In all other cases, the hierarchy of generality is ambiguous.

Sometimes, generality is associated with simplicity. This would be the case if model S had extra conditions of application as well as all of the conditions of application of model G. An example of simplification implying generalisation will be explored in the companion to this paper (part 2).

2.3 Truth

Truth is a difficult philosophical concept that will *not* be addressed in this paper. Instead, we take a more pragmatic approach relating to model purpose and validation.

Social scientists following the methodological position of Popper or Friedman essentially identify the truth of a theory or model with its predictive power. Nuanced differences in their positions are not germane to the present discussion. Experience shows that social science models do not have predictive power whether in the short or the long run.

Other social scientists adopt a post-modernist stance which, following Derrida, asserts that any text has to be interpreted and subject to criticism entirely on its own terms. In social scientific terms this comes down to something like "reality is socially constructed" so that one cannot evaluate any social analysis without knowing the social basis of the analyst.

Evidence-driven agent based modelling with expert/stakeholder participation falls in neither of these methodological camps. The purpose of the modelling is to capture the perceptions and self-described behaviour of the stakeholders in a way that formalises their expressions of those perceptions and descriptions. The validation the models capturing those perceptions and descriptions is achieved by

- implementing agent processing as a rule based implementation in which the rules are logic-like if-then statements couched in the linguistic terms used by the stakeholders and experts themselves
- presenting suitable formulations of the rules to stakeholder/experts for confirmation that the rules suitably capture their views
- implement the models to produce traces of agent behaviour and social interactions together with descriptions of the reasons for such behaviour and interactions
- present suitable formulations of the traces to the stakeholders and experts for confirmation of their plausibility, implausibility or simply their unexpected features.

Evidently, the purpose of the model is not to provide some objective or external truth but to capture actors' and experts' perceptions in ways that they find appropriate and enlightening.

This statement of "truth" implies an appropriate constraint on simplicity – but it is a constraint on maximal complication rather than maximal simplicity. In relation to the *obiter dictum* of William of Ockham, the number of entities in the model is determined by evidence and available computational resources. The modeller specifies the types of individuals and such groupings as households and churches. Agents are designed to represent individuals in these groupings. The number of agents is the lesser of the size of the population being modeled and the number of agents that can be implemented within the processing and memory constraints of the hardware and the time constraints of the modeller or model user.

3 General, true and simple social theory?

A social theory which is generally applicable, true in the sense that it describes observed social processes and outcomes and simple is unquestionably desirable. It is equally unquestionable that social science methodology has not successfully produced any such theory.

An alternative methodology based on the emerging approach to empirical, agent based, social simulation modelling starts from qualitative, narrative (some would say, anecdotal) evidence which is used to build detailed, empirical models. The models must be simple enough for the providers of the narratives – stakeholders and other domain experts – to be able to assess their plausibility and accuracy. A step towards simplicity and generality without losing accuracy can always be attempted by abstracting from the detail of the empirical model in a way which, for example, which substitutes abstract instantiations of classes of entities for empirically driven representations of entities. In the example of the companion paper to this, the modeller can substitute a set of abstract activities (`activity-1`, `activity-2`, ...) for specific churches, clubs, schools, *etc.* Then rules determining agent behaviour that depend on specific activities must be combined to specify behaviour activities in general. This step may involve fewer rules if some real activities require or motivate specialised behaviour. And model initialisation is likely to entail fewer lines of programming code to set up the activity entities since all that is required is a number of iterative loops equal to the number of activities being instantiated. There is nothing special about any of these instantiations. In such cases, both enhanced generality and simplicity result. The question of accuracy (or truth) is a different issue.

There is no reason to suppose that the informants who provided the narrative and qualitative accounts will be able to assess models at the resulting level of abstraction. However, if the empirical models have been validated by stakeholders and other domain experts and the abstract models yield the same results such as friendship networks with the same distribution of node degrees, the same patterns in the time series of cluster coefficients and the same time pattern of connectivity, then we can have some confidence (at least we have not disconfirmed) that the empirical model is a special case of a more general model. The model is more general in the sense that, at the very least, its assumptions describe conditions in which there are ranges of activities but not anything very specific about those activities. So the validated abstract model is simpler and, in conditions of application, more general than the empirical model and might also be true. To determine whether it is both more general and true will be a matter for further exploration. If other empirical models can be abstracted in the same way to yield the same model, then it would seem reasonable to have greater confidence that the abstract model is simpler and more general than the empirical models and equally true.

It seems worth noting that the approach to generalisation and validation described here is an alternative to the model-to-model (M2M) comparisons advocated by Edmonds, Rouchier, Hales and others [refs] – though by no means a mutually exclusive alternative. The fact is that model replication, though useful

and often showing up the effects of highly special constructions, is not obviously less time consuming than evidence based modelling and may also require detailed interaction with stakeholders – in this case the modellers whose models are to be replicated [Axtell-Epstein-Axelrod-Cohen]. Experience shows that abstraction is easier and more straightforward than replication. If the abstract models are also simpler in the sense of having fewer behavioural rules and lines of code, then it seems likely that M2M comparisons among abstract models will be easier than replications of empirical models. Both M2M analysis and the sort of model abstraction suggested here and demonstrated in the companion paper give us information about models as formal artifacts. As a result, we know more about the sensitivities of the models to particular assumptions and special programming devices. This knowledge is important in forming our own assessments of how well and robustly our models capture the social processes (or stakeholder perceptions of the social processes) under consideration. The further advantage of model abstraction is that, as part of a coherent program of companion modelling, it might lead to a collection of *canonical* models [Moss, 2000] that instantiate to a wide range of empirical models. The result would be an effective body of theory embodied in general and relatively simple models the goodness or truth value of which could be assessed by their history of validation.

Bibliography

- Olivier Barreteau, François Bousquet, and Jean-Marie Attonaty. Role-playing games for opening the black box of multi-agent systems: method and lessons of its application to senegal river valley irrigated systems. *Journal of Artificial Societies and Social Simulation*, 4(2): <http://www.soc.surrey.ac.uk/JASSS/4/2/5.html>, 2001.
- N. Becu, P. Perez, A. Walker, O. Barreteau, and C. Le Page. Agent based simulation of a small catchment water management in northern Thailand: Description of the CATCHSCAPE model. *Ecological Modelling*, 170(2-3):319–331, 2003.
- Alfred D. Chandler. *The Visible Hand: The Managerial Revolution in American Business*. Harvard University Press, 1977.
- Thomas E. Downing, Scott Moss, and Claudia Pahl Wostl. Understanding climate policy using participatory agent based social simulation. In Scott Moss and Paul Davidsson, editors, *Multi Agent Based Social Simulation*, volume 1979 of *Lecture Notes in Artificial Intelligence*, pages 198–213. Springer Verlag, Berlin, 2000.
- Bruce Edmonds. Simplicity is not truth-indicative. In Carlos Gershenson, Diederik Aerts, and Bruce Edmonds, editors, *Philosophy and Complexity, Worldviews, Science and Us*, pages 65–80. World Scientific, Singapore, 2007.
- S. Moss. Critical incident management: An empirically derived computational model. *Journal of Artificial Societies and Social Simulation*, 1(4), 1998.
- Scott Moss. Canonical tasks, environments and models for social simulation. *Computational and Mathematical Organization Theory*, 6(3):249–275, 2000.
- Edith T. Penrose. *The Theory of the Growth of the Firm*. Basil Blackwell, Oxford, 1959.
- Nathan Rosenberg. *Perspectives on Technology*. Cambridge University Press, Cambridge, 1976.