

Policy analysis from first principles

Scott Moss*

Centre for Policy Modelling, Business School, Manchester Metropolitan University, Manchester M1 3GH, United Kingdom

The argument of this paper is predicated on the view that social science should start with observation and the specification of a problem to be solved. On that basis, the appropriate properties and conditions of application of relevant tools of analysis should be defined. Evidence is adduced from data for sales volumes and values of a disparate range of goods to show that frequency distributions are commonly fat-tailed. This result implies that any stable population distribution will generally have infinite variance and perhaps undefined mean. Models with agents that reason about their behavior and are influenced by, but do not imitate, other agents known to them will typically generate fat-tailed time series data. A simulation model of intermediated exchange is reported that is populated by such agents and yields the same type of fat-tailed time series and cross-sectional data that is found in data for fast moving consumer goods and for retail outlets. This result supports the proposition that adaptive agent models of markets with agents that reason and are socially embedded have the same statistical signatures as real markets. Whereas this statistical signature precludes any conventional hypothesis testing or forecasting, these models do offer unique opportunities for validation on the basis of domain expertise and qualitative data. Perhaps the most striking conclusion is that neither current social theory nor any similar construct will ever support an effective policy analysis. However, adaptive agent modeling is an effective substitute when embedded in a wider policy analysis procedure.

The Issues

The purpose of this colloquium is to explore adaptive agent models and, in so doing, to force reexamination of current social theory and encourage rethinking of the processes by which human organization emerges.

The presumption in the meeting overview, from which the above passage is extracted, is that adaptive agent models are particularly well suited to capture the nature and consequences of social complexity whereas current social theory is not. Two aspects of current social theory are identified in the meeting overview: the dominance of social theory based on the assumption that economic actors maximize utility and the assumption that “social organization evolves from a top-down hierarchical system of culture and norms that serves to shape individual behavior.” It is then asserted that “adaptive agents methods are likely to become the foundations for modeling and simulation that may help to resolve many of the problems of complexity and help in the development of policy tools that provide enhanced insight into the likely effects of policy action.”

Why should adaptive agents models be relevant to, much less force the reexamination of, current social theory? And why might adaptive agents models usefully inform the provision of enhanced insight into the likely effects of policy action? After all, agent-based simulation is hardly the first analytical approach for which great promise has been claimed. Game theory and then dynamic game theory were going to provide powerful and relevant models of competition. Yet a recent survey (1) of game theory papers in a leading economic journal showed that game theoretical analyses of processes were limited to two or in one case three agents and all n -person game theoretical models were

concerned with equilibrium outcomes rather than any process. Econometrics and Keynesian theory together were going to provide means of forecasting the effects of policy actions provided that the number of policy targets was the same as the number of instruments. I recently asked on the e-mail discussion list of the International Institute of Forecasters whether there are any counterexamples to the following claim: “Since the invention of econometrics by Jan Tinbergen in the 1930s, there has not been a single correct econometric forecast of an extreme event such as a turning point in a trade cycle or a stock market crash. Every such forecast—without exception—has yielded either a type I or a type II error.”

Apart from one undocumentable claim, the strongest responses were that, when applied to past data, some new modeling techniques look better than most previous modeling techniques. No one was able to point to a correct forecast in real time.

This experience and many like it across the social sciences are reasons enough carefully to investigate the claim that agents and simulation are indeed promising elements in a new approach to policy analysis. A careful investigation specifies the problem to be addressed and does not alter the problem specification to conform to the requirements of any tool of analysis. The selection of any analytical tool to be applied to the problem, or the requirements analysis of the properties of any such tool is to be based on available and applicable empirical evidence. The tool of analysis considered in this paper is broadly in the class of agent-based social simulation models.

The whole range of policy analysis in complex environments is much too broad to be the subject of the sort of careful investigation suggested here. It is, however, possible to address a class of policies: those that seek to use some market or competitive mechanism to manage resources. European privatizations of public utilities and transportation systems or the use of internal markets in the United Kingdom’s National Health Service or proposals for carbon trading and carbon taxes to mitigate the extent of anthropogenic global climate change are examples of such policies. The choice of this class of public policies is motivated both by the importance of the policy goals and by the importance of representations of markets in the development of much of the current social theory to which the meeting overview refers.

To evaluate policies intended to use or create markets, we require analytical tools that capture the properties of markets that the policies are intended to exploit. For this reason, we begin in section 2 by a consideration of data describing the demand for and sales of a range of goods. The result is a demonstration that even the most mundane goods are subject to the same sort of volatility and uncertainty that is found in financial market prices and sales volumes. Whereas the latter are usually ascribed to

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, “Adaptive Agents, Intelligence, and Emergent Human Organization: Capturing Complexity through Agent-Based Modeling,” held October 4–6, 2001, at the Arnold and Mabel Beckman Center of the National Academies of Science and Engineering in Irvine, CA.

Abbreviations: TVP, time varying parameters; SOC, self-organized criticality.

*E-mail: scott@cfpm.org.

speculative forces, it seems hard to argue that sales of shampoo or cookies are subject to speculative demands. There must be some deeper similarity. This finding motivates the discussion in section 3 of the relevant statistical issues from the standpoint of both econometrics and statistical physics. Three mutually exclusive possibilities consistent with the available data are canvassed, one of which is consistent with current social theory and two of which are consistent with adaptive agent models. In section 4, a model is reported to support the investigation of the implications of adaptive agent models for policy analysis. In section 5, we explore some wider issues concerning the use of adaptive agent models for policy analysis. Perhaps the most striking conclusion is that neither current social theory nor any similar construct will ever support an effective policy analysis. However, adaptive agent modeling is an effective substitute when embedded in a wider policy analysis procedure.

The Statistical Signatures of Competitive Intermediated Markets

If a competitive market is one where neither buyers nor sellers are able to set the prices in their transactions, then either pricing is the outcome of a process of negotiation or some third party must set the price. Such third parties are well known and include the market makers on the financial exchanges, retail shops, and processors and sellers of information, such as credit agencies. An important policy issue is whether competitive intermediated exchange is appropriate for allocating public services or the services of privately owned public utilities. If this is an appropriate arrangement, under what circumstances is it appropriate?

Economic propositions about the efficiency and social benefits of competition are based on equilibrium models. It is therefore worth asking whether equilibrium models actually provide the best available descriptions of the phenomena we observe. We begin with statistical observation.

Weekly scanner data from supermarkets show that sales of fast moving consumer goods such as alcoholic beverages in the United States and the United Kingdom and tea, biscuits, shaving preparations, and shampoo in the United Kingdom are marked by the kind of clustered volatility that we associate with asset prices in the financial markets.

Benoit Mandelbrot (2) first noted that log price changes in financial markets are commonly power law distributed. He pointed out that this phenomenon is consistent with a stable Paretian population distribution. The value of the stable Paretian distribution is that there is a known functional relationship between the moments of a distribution constructed by multiplying a constant by observations drawn from other distributions and the moments of those other distributions. These functional relationships underlie such important statistical techniques as regression and correlation analysis.

The characteristic function of the stable Paretian distribution takes the logarithmic form

$$\log \int_{-\infty}^{\infty} \exp(iuz) d \Pr(U < u) = i\delta z - \gamma |z|^{\alpha} \left[1 + i\beta \left(\frac{z}{|z|} \right) \tan \left(\frac{\alpha\pi}{2} \right) \right], \quad [1]$$

where α in the interval $[0,2]$ is a “peakedness parameter” and β in the interval $[-1,1]$ determines skewness. Together with values of δ and γ , these parameters determine the mean, variance, skewness, and kurtosis of the distribution.

The value of α is of the most interest here. When $\alpha = 2$, this characteristic function reduces to that of the normal distribution. For all values of $\alpha < 2$, the variance of the distribution is infinite. Moreover, for values of $\alpha < 1$, the mean of the distribution is undefined. This means that, for all values of $\alpha < 2$, the law of large numbers does not apply to the variance and, for all values

of $\alpha < 1$, the law of large numbers does not apply to the mean. No laws or theorems of classical statistics or econometrics are applicable in these circumstances.

It is, however, important to note that the central limit theorem will typically apply to data with a stable Paretian distribution. Because aggregating time series data—say daily into weekly data points—is effectively to calculate the mean of daily data over seven data points and then multiply by seven, samples of these sample distributions will be approximately normally distributed even if the underlying daily data are not. Consequently, it is possible to generate data that appear to be normally distributed simply by taking data of sufficiently low frequency. However, the variance and possibly the mean of the distribution of sample distributions will not converge to stable values.

Distributions with infinite variances are easily distinguished visually from normal distributions because, for the same means and variances, they have fatter tails and therefore thinner peaks—a condition known as *leptokurtosis*. Fig. 1 shows clear evidence of leptokurtosis in the weekly sales values of three brands of shampoo in United Kingdom supermarkets for the 65 weeks beginning 2nd January 2000 (Fig. 2). Similar results are found for virtually every one of the 120 or so brands of shampoo for which I have the data as well as every brands of tea, shaving preparations, biscuits, and, in the United States as well as the United Kingdom, every one of some 200 brands of spirituous alcoholic beverage and beers. The first row of Fig. 1 shows weekly sales values. Brand A is a leading brand with no discernible sales trend whereas sales values of brand B are declining and sales values of brand C are increasing. Both of the latter have small market shares. The second row shows the time series of relative sales changes. Over the 65 weeks, there were obvious clusters of volatility, and it is these clusters that generated the extreme values that cause the leptokurtosis evident in the third row, showing the frequency histograms of the relative sales changes compared with the corresponding normal distribution.

These results are typical for all of the products considered as well as daily metered consumption of water in southern England. Leptokurtosis and clustered volatility are evidently far more general than has previously been recognized. Without a doubt, the consequences for forecasting are enormous because leptokurtosis alone, independently of the clustering of volatility, implies the failure of the law of large numbers. That is, increasing sample size does not result in any convergence of any of the moments of the sample distributions—a result that itself renders parametric statistical forecasting techniques wholly otiose (3). Clearly, if leptokurtosis undercuts the law of large numbers, then the clusters of extreme events cannot in principle be forecast by statistical means.

Three Responses to Leptokurtosis and Clustered Volatility

There have been two responses to this problem, both of them addressing the core issue of the failure of the law of large numbers in samples of the variance of time series data. One of these responses has been offered by econometricians and is in effect intended to preserve the law of large numbers, and the other is from the physics community and is intended to bury the law of large numbers for good—at least when it comes to forecasting financial asset prices. A further response, drawing on agent-based social simulation, is described below.

Time Varying Parameters (TVP). The TVP approach is based on the assumption that the observed time series is drawn from a normally distributed population with constant mean and varying variance. The variance at the time of any observation is itself a function of previously observed errors so that standard regression techniques can be used to model the time series of variances. To capture the clustering of extreme events, the variance of a time series for any observation is treated as a function of

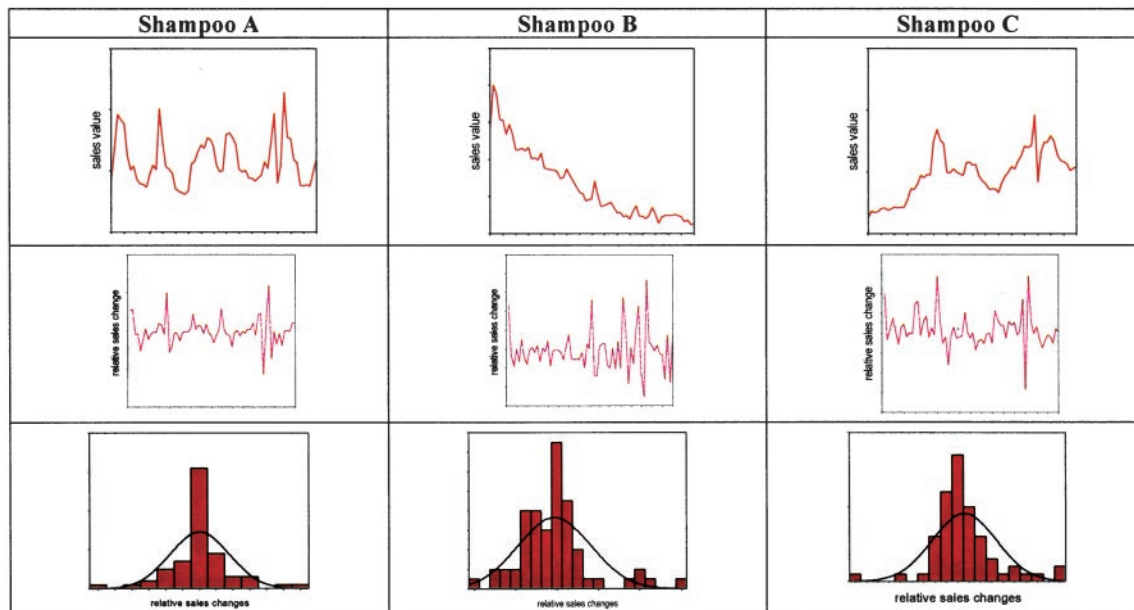


Fig. 1. Weekly shampoo sales and relative sales change: 2 January 2000–25 March 2001 (Source: Information Resources International).

previous error terms. Recent, large error terms tend to generate large variances. Because the observation is then drawn from a population with a larger variance, the probability of the observation being relatively far from the mean is greater than when the population variance is smaller.[†]

The motivations offered for particular TVP estimating methods are invariably related to rational expectations, the mean-variance representation of risk and risk aversion, or some similar equilibrium notion from economic theory. There are, however, no microeconomic equilibrium models that generate both leptokurtosis and clustered volatility either analytically or by means of simulation.

Self-Organized Criticality and Econophysics. Although there are no models demonstrating any microeconomic foundations of TVP-based leptokurtosis, there is an extensive class of models—both canonical and applied—that were developed and explored by statistical physicists because they do generate clustered volatility and hence leptokurtosis. The physical problem being addressed was the observation that an extraordinarily wide range of physical phenomena are power law distributed. The power law distribution is:

$$N(s) \sim s^{-\tau}, \quad [2]$$

where N is the number of observations at scale s and $\tau > 0$ is a parameter. Mandelbrot (2) pointed out that the power law distribution is a characteristic of the stable Paretian distribution.

The question of concern to statistical physicists starting with Per Bak and his colleagues (8) was to find the process that is both very general and yields power law-distributed time series. The canonical model they developed was an idealization of a sandpile, with grains of sand being continually added. The sandpile model is closely related to a cellular automaton model in that it is located on a grid with nonperiodic boundaries with grains of sand added to cells at each time step. Whenever the number of grains of sand in a cell reaches some specified critical level—say four—there is a “toppling” of the sand in that cell. This toppling

takes the form of a redistribution of the grains of sand in the critical cell to other (not necessarily adjacent) cells in the grid. Not all of the grains are redistributed, but the number of grains in the critical cell is nonetheless reduced to 0. That some grains are lost from the system in this way makes it dissipative.

Of course, adding toppled sand to the grains at other cells increases the numbers in those cells until some of them become critical and topple and so increase the number of grains in yet other cells, and so on. The consequence is that, once the system reaches a critical state, there will be a sequence of topplings involving different numbers of cells in the grid. The time series of these topplings is power law distributed.

There is a growing family of such models that yield power law distributed time series and cross-sectional data. The key feature of these models is that they do not require fine tuning of the parameters of interest to produce data with this statistical signature. In this sense, the models self-organize into the critical state and remain in that state thereby to produce such power law-distributed data with clusters of extreme events.

Jensen (9) has summarized the conditions in which self-organized criticality (SOC) emerges as follows.

- Model components (cells, agents, etc.) are metastable in the sense that they do not change their behavior until some level of stimulus has been reached.
- Interaction among the model components is a dominant feature of the model dynamics.
- The model is a dissipative system.
- The system is slowly driven so that most components are below their threshold (or critical) states most of the time.

In social terms, agents and the individuals they represent are metastable if they do not respond to every minute stimulus they face. They would not, for example, reconfigure their desired shopping basket as a result of a penny rise in the price of a tin of tuna. A particular implication of metastability is that the behavior of individuals cannot be represented by utility maximizing software agents. The dominance of interaction among the agents amounts to social embeddedness in the sense of Granovetter (10) and Edmonds (11): the behavior of individuals cannot be explained except in terms of their interaction with other individuals known to them. Dissipation in a social system,

[†]Bollerslev (4) identifies the core econometric processes of relevance here to be the ARCH process [Engle (5)], the GMM process [Hansen (6)], and GARCH [Bollerslev (7)].

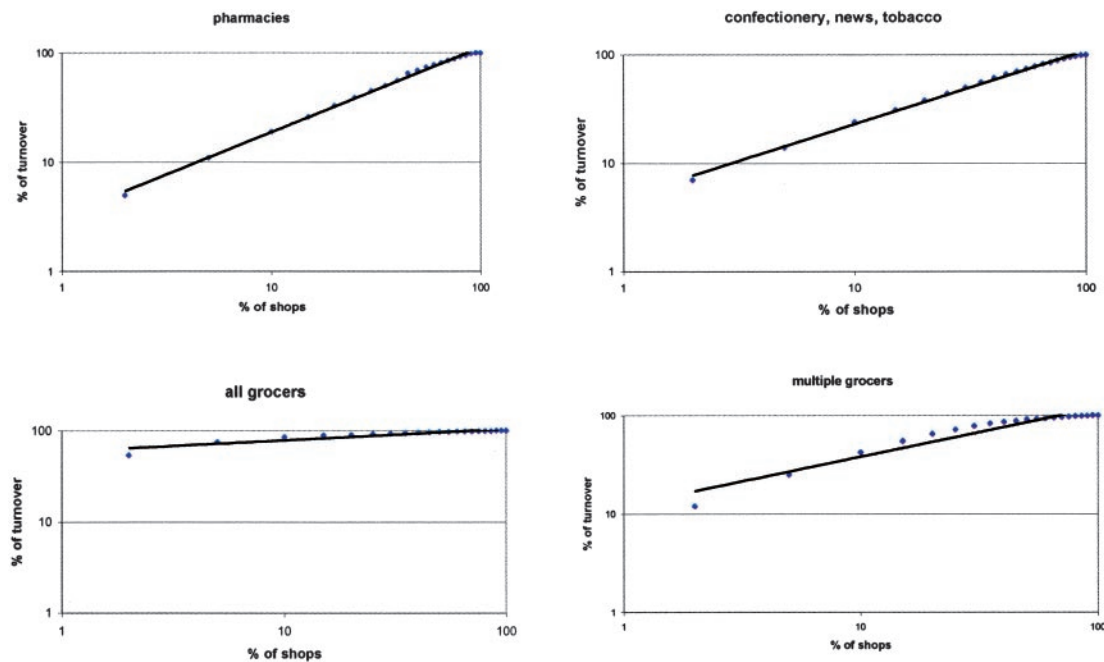


Fig. 2. Market share distribution of United Kingdom retail outlets (19).

analogous to the dissipation of grains of sand in the sandpile model, equates to individuals being influenced by other individuals without slavishly imitating them.

There is a literature on SOC models of financial markets although these articles appear almost entirely in journals such as *Nature*, *Physica A*, *Physical Review E*, and *Physical Review Letters* and, apart perhaps from Lux (12), have had no noticeable effect on the economics literature.

There are several important differences between the self-organized criticality and the TVP literatures. The former uses models to generate statistical signatures that do not replicate actual data series whereas the latter exploits the data to model changing values of the moments of a distribution function. This is part and parcel of a more fundamental difference: self-organized criticality suggests that, without unlimited computational and information processing capacities, forecasting extreme events is inherently infeasible whereas the TVP literature is based on a faith (i.e., without supporting evidence) that forecasting extreme events is feasible. A further difference is that SOC properties are, with a few special exceptions, known only from simulation experiments. There are hardly any analytical proofs (9). The TVP literature, by contrast, is based on algorithms with properties that have been proved analytically.

SOC in Social Systems. One question that has not been considered in either the TVP or the SOC literature is whether there is any population distribution underlying observed or simulated time series or cross-sectional data. Because the TVP literature is concerned with the application of parametric statistical techniques, the assumption of a population distribution with fixed characteristics is essential. In the TVP literature, there is assumed to be a stable distribution relating variance to previous deviations from a fixed mean. In the physics literature on SOC, a measure of success is naturally taken to be the degree of agreement between the power law distribution parameter obtained by simulation and the parameter obtained from the corresponding real data. Whereas this measure of success is appropriate in models of physical systems, it may not be an appropriate measure of success of models of social systems.

The key difference here turns on universality. The assumption that the laws of physics are always and everywhere the same has been enormously useful in the physical sciences. It is also a natural assumption to make because fundamental physical relations are plausibly unchanged by their own consequences. For example, the law of gravity does not vary because objects catastrophically collide. It is much less plausible to argue that social relations are unchanged by their consequences. On a grand scale, it would be lunatic to suggest that social relations were unchanged by the French Revolution—and not just in France. More prosaically, major financial panics typically result in changes in the rules and practices of financial markets. Also, institutional arrangements are altered by extreme natural events, but the laws of nature are not affected by extreme social events.

If we consider the natural and social systems as two data-generating mechanisms, then we observe that the data emerging from natural systems typically have fixed statistical properties whereas data emerging from social systems typically do not. For example, earthquake magnitude distributions obey the Gutenberg-Richter law, and the observed distribution does not change over time. So, whereas it is not possible to predict the occurrence of specific earthquakes or earthquakes of specific magnitudes, it is possible to describe with considerable accuracy the distribution of earthquake magnitudes. In social systems, however, experience of extreme events leads to a search for means of reducing their incidence if that is possible or their impact if it is not. The point is to change the observed distribution by, in effect, changing the data-generating mechanism. To the extent that such social engineering is successful, the parameter of the power law distribution will be reduced although the reduced goodness of fit will not eliminate the leptokurtosis and clustered volatility of the observed data.

Implications for Scientific Method

Three mutually exclusive explanations of observed leptokurtic data series were described above: a normal distribution with predictably time-varying parameters, a stable Paretian distribution with infinite variance generated by a self-organized critical social process, and data generated by a self-organized critical

social process but not behaving like a sample of any fixed population distribution. There are no tests on observed data that will distinguish between TVP, stable Paretian distributions, or leptokurtic observations not drawn from a population distribution. Clearly, a systematic history of success at forecasting volatile episodes with TVP methods would give convincing support to the hypothesis that there is an underlying normal distribution with a predictably varying variance. However, there is no such history. The timing, magnitude, and duration of volatile episodes remain in practice unpredictable. We must look to some other means of discriminating among—or rejecting—these possibilities.

In the natural sciences, the search for explanations of previously unexplained observations has taken the form of a search for a data-generating mechanism that could be validated independently of the observations themselves. A classic example is the validation of general relativity theory by comparing observations of star positions during a total solar eclipse with the predictions of the theory. A more pertinent example is the development of the sandpile model to explain observed power law distributions and then experimental testing of the canonical model with sand, rice grains, and the like (13).

Scientific methods that have proved to be successful in the natural sciences are not necessarily equally applicable in the social sciences. However, it is hard to see any objection to treating a social system as a data-generating mechanism and devising a model to represent that mechanism. If the model captures self-organized criticality, then it is not possible to validate it by statistical means for two reasons. One, of course, is that the generated data will be leptokurtic and therefore have infinite variance. There are no parametric hypothesis testing procedures for infinite variance distributions, and nonparametric procedures provide information only about the data in hand. The second reason is that self-organized criticality implies that the timing, magnitude, and duration of clusters of extreme events are in practice unpredictable. There seem to be two, mutually exclusive ways forward.

One is to continue to develop the TVP approach in the hope that it will someday yield consistently accurate forecasts of extreme events. The other is to focus on the relevant system as a data-generating mechanism and to devise means of modeling that mechanism *independently* of the statistical data itself.

From Problem to Approach. The issue of concern is policy analysis in conditions where the objective is to develop strategies to mitigate the impacts of clusters of extreme events, the magnitude, duration, and timing of which cannot be forecast. The suggested alternative to forecasting is to try to understand the social processes generating the extreme event clusters to assess the effectiveness of different responses to their occurrence. The means to be chosen for understanding the underlying data generating process must obviously be able to capture a process yielding unpredictable extreme event clusters. This requirement filters out all equilibrium-generating processes.

A second requirement is that the approach should be robust in explaining leptokurtosis and unpredictable clusters of extreme events. A mathematical model that generates data with the appropriate statistical signature only under a very narrow range of values of key parameters would not be appropriate unless there was independent evidence that those parameters and the particular values required robustly describe observed phenomena. Whereas chaos and edge of chaos models meet the first requirement, there is some evidence that they do not meet the second. Differential or difference equation models with strange attractors have been known from the discovery of chaos to require parameters to be set to specific ranges. Kaufman (14) has worked on models in which values of those parameters are driven to the chaotic range and are otherwise at “the edge of chaos.”

However, Per Bak (13) reports that these results are not robust with respect to parameter settings and initial conditions. If Bak is wrong, then edge of chaos approaches will satisfy the second requirement. If he is right, they will not.

The third requirement is that the approach must support independent validation. To date, there has been no independent validation of the edge of chaos models.

The approach investigated here starts expositionally from SOC. Historically, however, a set of models designed and implemented in ignorance of SOC by researchers in the Centre for Policy Modeling to analyze the effects of social embeddedness and to be open to validation by stakeholders turned out to yield leptokurtic data with clustered volatility.

Two examples of such models are Moss’s model of household water demand and the effects of exhortation by government and other authorities during conditions of drought (15) and Edmonds’ model of a financial market. Moss represented agent cognition by means of a combination of the problem space architecture of Soar (16) and ACT-R (17) together with an endorsement mechanism (18). Edmonds represented agent cognition by means of an elaborated genetic programming algorithm. Both of these representations of cognition yield metastable agent behavior in that some nonnegligible weight of evidence and incentive is required to induce agents to change their behavior. In both models, agents were socially embedded. In the Moss model, social embeddedness took the form of observation of neighbors’ public consumption activities such as garden watering and car washing as well as word of mouth communication. In the Edmonds model, there is word of mouth communication among agents. In both cases, agents were influenced in their behavior by the behavior of, and communication with, the subset of other agents with which they had formed some relationship of trust and regard. This result, of course, is the essence of social embeddedness.

In these social simulation models, the representations of agent cognition could hardly have been more different. Even the representations of the model spaces were different—Moss implemented a grid with periodic boundaries whereas Edmonds had no spatial location of agents. Indeed, apart from metastability of agent behavior and social embeddedness, it is hard to identify common features in the design and implementation of the two models. The natural conclusion is that, on the basis of such experience, cognitive representations yielding metastable agent behavior and social embeddedness of agents drive processes yielding leptokurtic distributions with clustered episodes of volatility.

It would be wrong to expect such social simulation models to replicate observed time series from their target systems. What is being sought in these model specifications is a shared statistical signature in the sense that the time series of both model and target systems are marked by unpredictable clustered volatility and therefore leptokurtosis. If volatility is unpredictable in both systems, we can hardly expect to be able to engineer the model system to replicate the time pattern of volatility of the target system because, to do so would render the volatility of the target system predictable!

It is nonetheless possible to validate the goodness of the representation of the target system by the model system. The validation must be more direct and expressive than statistical validation techniques. The basis of the validation technique is the implementation of agents to represent specific observable social entities. Such entities could be individual persons or collections of persons as constituent components (departments, sections, or the like) of organizations, whole organizations, government agencies, or any other recognizable entity. Validation must be undertaken in collaboration with (or actually by) domain experts who know the behavior of the target entities and the ways they interact with other social entities. The key question here is

whether such behavior is plausible to the domain experts who may themselves be the target entities or members of them. In this case, the domain experts are participating stakeholders.

Because every model is sensitive to the values of some parameters, an essential element of model validation is that either the model behavior is not sensitive to parameter values without unobservable target statistics or the model endogenously drives key values into the range that supports the replication of the behavior of the target system.

An Example System. In addition to the finding reported above that sales volumes and values by brand have the indicated statistical signature, market shares by retail outlet are power law distributed when the markets are competitive. The data for three branches of the retail trades are reproduced in Fig. 2. The suggestion that market shares might not be found for the less competitive trades is due to the deviation of the market shares from the power law distribution for multiple United Kingdom grocers but not (or at least much less markedly) for all grocers. In light of these data, we would expect a model of a competitive market with intermediaries to yield power law-distributed market shares.

This proposition is tested with a model of a market in which there are adaptive agents representing customers and adaptive agents representing intermediaries.

There are also product sources that are not given any representation of cognition. The social network in this model is represented by a grid with periodic boundaries in which agents can “see” a limited number of cells in each of the four cardinal directions.

Cognitive agents in the model buy and/or sell items represented by the values of digits in an ordered list—a digit string. The values of the digits in the string can be to any arbitrary base. At each trading cycle, a digit string generator produces a digit string. The length of the string is constant over each simulation run.

There is a user-determined number of product sources distributed at random on the grid. Each source holds the current values of digits at specified positions in the digit string. These values change as the system digit string changes.

The intermediaries acquire the values of digits from sources. These values can be acquired only as packets of all items held by a source. However, the intermediaries can sell items individually or in any combinations available to them, selling on to other agents only those items the other agents demand. Moreover, intermediaries can combine the items acquired from several sources. There is a flow of intermediaries chosen at random from the $[1, B]$ interval where B is the maximum number of intermediaries, set by the model operator, that can enter the market at each trading cycle. Each intermediary builds asset reserves from profits and leaves the market when its asset reserves are exhausted.

Each intermediary is initially allocated to an empty cell but can choose to move to some other cell if it is unoccupied and no other agent is seeking to move at the same time to the same cell. The motivation to change cells is the knowledge that there is a profitable intermediary in the neighborhood of the destination cell.

Either customers acquire packets of items from sources in the same way as do the intermediaries or they buy just the items they want from the intermediaries. The customer agents each inhabit a cell during the whole of the simulation run. Although the number of customer agents is determined at the start of each run by the model operator, their locations are determined at random.

At the start of each simulation run, customer agents are allocated demands for the values of digits at specified positions in the system digit string. The number of items demanded is determined at random in an interval set by the model operator

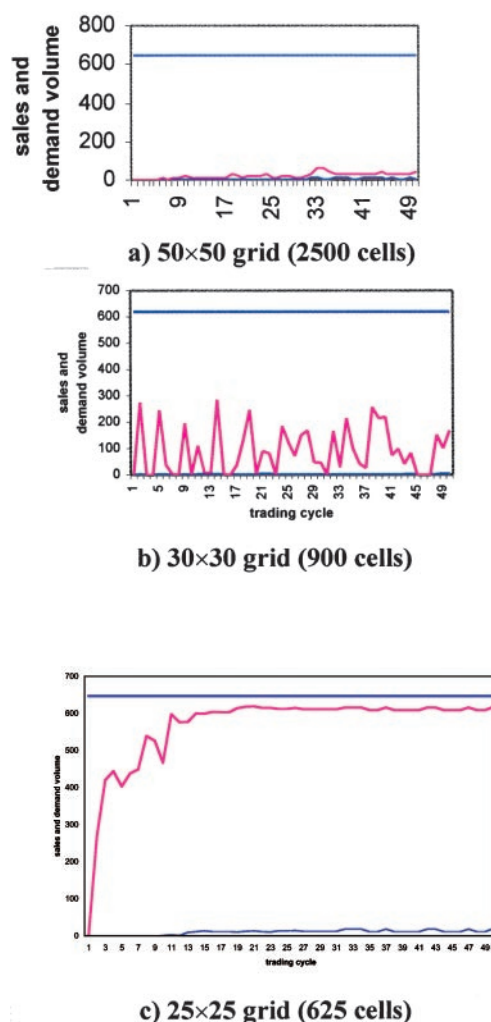


Fig. 3. Sales volumes and demands at different agent densities.

at the start of the simulation run. Intermediaries demand only items for which they have previously received inquiries from customers or other intermediaries.

Intermediaries and customers are synchronous, parallel agents. To enable them to communicate with one another, a series of *communication cycles* is nested within each trading cycle.

In all of the simulation runs, the system digit string contained 40 digits; there were 15 sources and 100 customers. Each customer could demand up to 12 items, and each source could hold up to 15 items. The maximum number of broker agents entering the market in any trading cycle was 15. Agents could identify the existence of sources or other agents within eight cells of their own position in the cardinal directions (up, down, right, and left). The only parameter setting that was changed for the different simulation runs was the size of the grid. Three grid sizes were used: 50×50 (2500 cells), 30×30 (900 cells), and 25×25 (625 cells). A larger grid size implies a lower density of agents.

Experimentation confirmed that agent density is a critical factor in the viability of agent trading, that a high proportion of demands are satisfied only when virtually all trading is intermediated and market shares are leptokurtic.

A natural measure of the effectiveness of markets is the proportion of total customer demands that are satisfied through transactions. The time series of these proportions for three scales of grids are shown in Fig. 3. The population density of customers

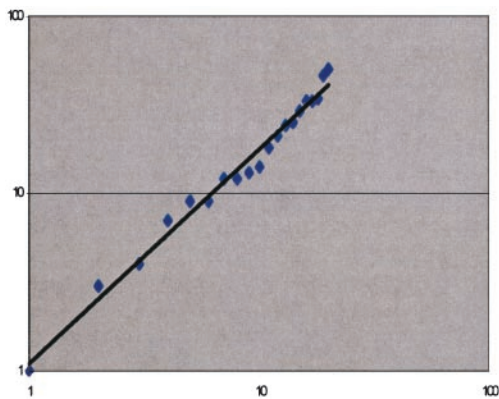


Fig. 4. Intermediaries' market share distribution at the 50th trading cycle.

and sources increases from a down to c , with the corresponding proportion of satisfied demands rising from 3.2 to 14.6 to well over 90%.

Demand satisfaction in all of the modeled markets was a result of intermediated transactions. In Fig. 3, the lower line in each case represents acquisitions of items by customers directly from sources. Evidently, in all cases, direct acquisition from sources was negligible.

The statistical signature identified for real intermediated markets is replicated by the simulation model. Fig. 4 shows that the power law holds for cumulative sales volume against the rank of the broker (from lowest to highest sales) at the 50th trading cycle of the simulation of the 625-cell market. Although the power law distribution prevailed consistently during all trading cycles, the parameters of that distribution were changing over time. The frequency distribution in Fig. 5 demonstrates that leptokurtosis of each intermediary's sales volume changes mirrors that of brand data as reported above.

Adaptive Agent Models and the Process of Policy Analysis

The model reported above has been validated to the extent that its statistical signature shares leptokurtosis, clustered episodes of volatility, and power law-distributed market shares with high frequency data captured from supermarket and other retail sales outlets. This result coheres with the longstanding results on high frequency price and volume data from organized financial markets. The differences in the uses to which items traded in these markets are put suggest that the common factor among these markets considered as data-generating mechanisms is that

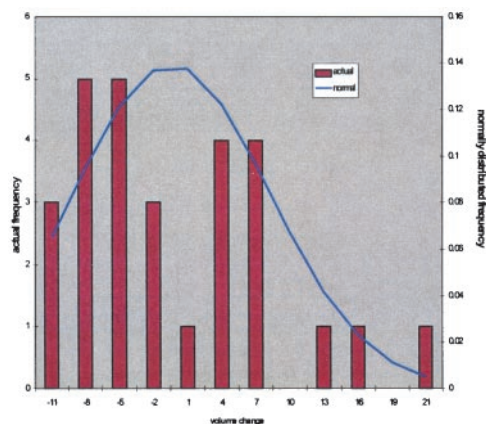


Fig. 5. Frequency distribution of simulated sales volume changes—actual and normal.

they are competitive and that transactions take place through intermediaries—brokers, supermarkets, newsagents, etc. The need for competitiveness is consistent with the implication from the simulation experiments that some critical density of agents is required for exchange to be conducted efficiently.

Whereas this result supports a general presumption about the requirements for efficient trading in intermediated markets, it would be rash to build any policy prescriptions thereupon. For one implication of the validation process reported here is that the models cannot in principle be used for purposes of the prediction of events defined by the time of their occurrence and their magnitude—including the prediction of specific outcomes from any policy actions. Indeed, if SOC models support accurate descriptions of social systems as data-generating mechanisms, then the sort of prediction sought by physical scientists are *in principle* impossible to achieve in those social systems. Consequently, the use of adaptive agent models where the agents are metastable and their behavior is influenced by the behavior of other agents is strictly incompatible with the positivist approaches that justify current social theory.

In general terms, the approach of positivist social scientists is to start from a social theory, derive a specific model from that theory (usually a regression model), and then to apply that model to the data. The approach taken in this paper has been to identify the statistical signature of the data and then to consider alternative means of capturing those data. The advantage of the adaptive agent models is that they can (and the statistical models cannot) be used to describe components of the data-generating mechanism in arbitrary detail. Consequently, these models can be treated as descriptions of the data-generating mechanism. The validation of these models entails an assessment of the accuracy of those descriptions as well as an assessment of the accuracy of their statistical signatures.

The model reported here captures the leptokurtosis and clustered volatility of the relevant empirical data *and* captures the behavior and interactions of the relevant social actors. This result leaves two questions: (i) What does it mean to “capture” individual behavior and interactions? (ii) How can the models validated in this way be used for policy analysis?

Policy analysis by its nature targets existing social systems. Consequently, there must be stakeholders who are sources of expert information concerning the particular social domains of concern to the policy analysts. Stakeholders and independent domain experts can provide descriptions of the goals and actions of the relevant actors as well as the patterns and modes of interaction among them. They can also evaluate the plausibility of the models designed to incorporate those descriptions in the software code that constitutes each agent. A good agent-based model for these purposes will provide information about the agents' goals and behavior in a form that will enable stakeholders and independent domain experts to evaluate that behavior as descriptions of actual social entities. The “capture” of individual behavior and interaction is the design of agents and interaction mechanisms that define software systems (models) that generate system data with the appropriate statistical signatures and produce data about the agents and mechanisms that are validated as accurate or plausible by domain experts.

Stakeholder participation entails not only validation by domain experts but also a more organic process of development of the models in which the stakeholders both explicate and refine their understanding of the target systems and use the models to investigate alternative policy or other strategic options. In the latter case, the models are being used in the same sense as flight simulators to develop responses and abilities to identify the emergence of critical events at a relatively early stage. There is of course an important difference from flight simulators. Flight simulators are based on systems that are well understood and based on clear principals of good physical science and engineer-

ing practice that support clear predictions of the outcomes of actions affecting the system. We have seen that systems well described by SOC models are not well understood in this sense. Consequently, stakeholder participation in the modeling process requires that models be based on good science where that is useful but that stakeholder perceptions take precedence in model design over any conceptual frameworks or system representations developed by independent (e.g., academic) observers.

An important prospect here, currently being explored in the European Union-funded project on Freshwater Integrated Resource Management with Agents (FIRMA; <http://www.cpm.mmu.ac.uk/firma>) is the development of models by different stakeholders for use in, as one example, negotiation regarding measures for flood safety, water quality, environmental preservation and development, navigation, and economic exploitation of the Limberg basin of the River Meuse in The Netherlands. The stakeholders include a ministry of central government, the provincial government of Limberg, non-governmental organizations concerned with environmental issues, farmers, community groups, commercial companies, and an organization established to coordinate these various, conflicting interests. The process of participatory agent-based social

simulation modeling is used to identify conflicts in goals and in perceptions of existing conditions and the consequences of alternative courses of action. It is intended to graft segments of models developed with one set of stakeholders onto models representing the understanding of other stakeholders to clarify differences and to provide each stakeholder with a greater understanding of the interests and concerns of the other stakeholders.

This sort of process is very different from conventional investigation in the social sciences. Instead of developing a particular model based on some more general construct (theory), the models are devised on the basis of observation and developed by means of a process of empirical validation. Particularly for models incorporating SOC, no hypothesis testing procedures from classical statistics is appropriate and no predictions of particular events are supported. It may be that model development with stakeholder participation will lead to some more general propositions that can inform social or physical or biological theory. However, the usefulness of agent-based social simulation models developed with stakeholder participation is that they support the development of a social process of policy and strategic analysis when forecasting and prediction is infeasible with respect to the relevant natural and social systems.

1. Moss, S. (2001) *J. Artif. Soc. Soc. Simul.* **4**, <http://jasss.soc.surrey.ac.uk/4/2/2.html>.
2. Mandelbrot, B. (1963) *J. Bus.* **36**, 394–419.
3. Mandelbrot, B. (1997) *Fractales, Hasard et Finance* (Flammarion, Paris).
4. Bollerslev, T. (2001) *J. Econometr.* **100**, 41–51.
5. Engle, R. F. (1982) *Econometrica* **50**, 987–1007.
6. Hansen, L. P. (1982) *Econometrica* **50**, 1029–1054.
7. Bollerslev, T. (1986) *J. Econometr.* **31**, 307–327.
8. Bak, P., Tang, C. & Wiesenfeld, K. (1987) *Phys. Rev. Lett.* **59**, 381–384.
9. Jensen, H. (1998) *Self-Organized Criticality: Emergent Complex Behavior in Physical and Biological Systems* (Cambridge Univ. Press, Cambridge, U.K.).
10. Granovetter, M. (1985) *Am. J. Sociol.* **91**, 481–510.
11. Edmonds, B. (1999) *Adapt. Behav.* **7**, 323–348.
12. Lux, T. (1998) *J. Econ. Behav. Org.* **33**, 143–165.
13. Bak, P. (1997) *How Nature Works: The Science of Self Organized Criticality* (Oxford Univ. Press, Oxford).
14. Kaufman, S. A. (1993) *The Origins of Order* (Oxford Univ. Press, New York).
15. Downing, T. E., Moss, S. & Pahl Wostl, C. (2000) in *Multi Agent Based Social Simulation*, eds. Moss, S. & Davidsson, P. (Springer, Berlin), Vol. 1979, pp. 198–213.
16. Laird, J. E., Newell, A. & Rosenbloom, P. S. (1987) *Artif. Intell.* **33**, 1–64.
17. Anderson, J. R. (1993) *Rules of the Mind* (Lawrence Erlbaum Associates, Hillsdale, NJ).
18. Cohen, P. R. (1985) *Heuristic Reasoning: An Artificial Intelligence Approach* (Pitman Advanced Publishing Program, Boston).
19. A. C. Neilsen & Co. (1992) *The Retail Pocket Book 1993* (NTC Publications Ltd., Henley-on-Thames, U.K.).