STRATEGIC ORGANIZATION Vol 3(2): 219–228
DOI: 10.1177/1476127005052700
Copyright ©2005 Sage Publications (London, Thousand Oaks, CA and New Delhi)
http://soq.sagepub.com

SO!APBOX
EDITORIAL ESSAY

Why Gaussian statistics are mostly wrong for strategic organization

Bill McKelvey The UCLA Anderson School of Management, USA

Pierpaolo Andriani Durham Business School, UK

Organizational researchers presume Gaussian (normal) distributions, with stable means and finite variances, with appropriate statistics to match. For evidence, study any random sample of current quantitative research papers. It follows that nearly all of our research-based lessons to managers stem from Gaussian-based research. Suppose this premise is mostly wrong. What then?

Power laws

Consider the coasts of England and Norway. They appear jagged no matter what kind of measure is used: miles, kilometers, meters, centimeters, millimeters. This is called 'scalability': no matter what the scale of measurement, the phenomenon appears the same. Scalability results from what Benoit Mandelbrot (1983) called 'fractal geometry'. A cauliflower is an obvious example. Cut off a branch; cut a smaller branch from the first branch; then an even smaller one; and then even another, etc. Now set them all on a table, in line. Each fractal subcomponent is smaller than the former; each has the same shape and structure. They exhibit a 'power law effect' because they shrink by a fixed ratio.

Power laws underlie fractal geometry. They are 'indicative of correlated, cooperative phenomena between groups of interacting agents at the microscopic level' (Cook et al., 2004). They often take the form of rank/size expressions such as $F \sim N^{-\beta}$, where F is frequency, N is rank (the variable) and β, the exponent, is constant. In most exponential equations the exponent is a variable. Power laws call for 'scale-free theories' because the same theory applies to each of the different levels, that is, the explanation of the generative process is the same across all levels of analysis. The cauliflower is an obvious example of this as well.

Many complex systems resulting from emergent dynamics tend to be 'self-similar' across levels. The same process drives order-creation behaviors across multiple levels of an emergent system (Kaye, 1993; Casti, 1994; West et al., 1997). These processes are called 'scaling laws' because they represent empirically discovered system attributes applying similarly across many orders of magnitude (Zipf, 1949). Brock (2000: 30) observes that the study of complexity 'tries to understand the forces that underlie the patterns or scaling laws that develop' as newly ordered systems emerge. While scale-free theory is something else that is missing in strategic organization, we only deal with the statistical consequences here.

Power law phenomena exhibit Paretian rather than Gaussian distributions. The fundamental difference lies in assumptions about the correlation among events. In a Gaussian distribution events are assumed to be *in*dependent. Independent events generate normal distributions, which sit at the heart of modern statistics. When events are *inter*dependent, normality in distributions is not the norm. Instead Paretian distributions dominate because extreme events occur more frequently than normal, bell-shaped Gaussian-based statistics would lead us to expect. Physical, biological, ecological, social and industrial systems show an impressive variety of fractals (Kaye, 1993); and we list some in Table 1. Many scholars now believe that power laws are the best analytical framework to describe the origin and shape of most natural objects. Given the ubiquity of these discoveries and the nature of the underlying scale-free theory, we think they are unknown, unappreciated, but equally ubiquitous phenomena in organizations (Andriani, 2003).

Extremes versus averages

Linear thinking is engrained in our mentality. Scientific and mathematical models are based on the concepts of equilibrium and linearity. Linearity means two things: proportionality between cause and effect; and superposition, that is, when the dynamics of a system can be reconstructed by summing up the effects of the single causes acting on the single components (Nicolis and Prigogine, 1989), which allows efficient causality to operate, equations to be solved and forecasting models elaborated. Economics, for instance, is almost theistic in its assumption that economic phenomena trend toward equilibrium (Mirowski, 1989). However, this assumption allows linear equations and analytical simplicity.

By focusing on systems in equilibrium, statisticians implicitly accept that the number of possible states that a system can attain is limited (and computable) and that the search time following the onset of instability (i.e. an exogenous shock) is short compared with the equilibrium time. For this to be true, the many elements comprising a system must be assumed independent data points; otherwise we could have interdependence, possible mutual causality and the occurrence of possible extreme event. If we take 100 companies of

Table 1 Listing of some power law discoveries*

		Natural science	
Cities	Traffic jams	Coastlines	Brush-fire damage
Water levels in the Nile	Hurricanes & floods	Earthquakes	Asteroid hits
Sun spots	Galactic structure	Sandpile avalanches	Brownian motion
Bach's music	Epidemics	Genetic circuitry	Metabolism of cells
Networks in the brain	Tumor growth	Biodiversity	Circulation in plants, animals
Langton's Game of Life	Fractals	Punctuated equilibrium	Mass extinction, explosions
Brain functioning	Predicting premature births	Laser technology evolution	Fractures of materials
		Social science	
Language	Social networks	Internet	Blockbuster drugs
Sexual conquests	'Fordist' power	Wealth	Citations
Co-authorships	Actor networks	Job vacancies	Salaries
Firm size	Supply-chains	Growth rates of firms	Growth & internal structure
Casualties in war	Countries' GDP growth rates	Stock price movements	Delinquency rates
Movie profits	Consumer product sales	Size of villages	Cotton prices
Economic fluctuations	Biotech alliance networks	Entrepreneurship/innovation	

*Sources are given in Andriani and McKelvey (2005).

approximately of the same size and belonging to the same sector, assume independence and plot a variable, say profit, we discover that most events will pack around the average, in a rapidly decaying distribution that follows a bell curve. This distribution is by far the most studied statistical distribution; it is assumed to correctly characterize most of our discoveries about the natural and social worlds. *However, the crux of the point is independence of events.* In real life, these companies could: benchmark against each other, imitate those perceived as successful, exchange information, organize cartels, pursue mergers and acquisitions, compete for limited resources, etc. In a word, they are *inter*dependent and not *in*dependent. The statistical distribution governing interconnected agents does not give rise to a bell distribution but instead to a power law – a Paretian – distribution.

Gaussian and Paretian distributions differ radically. The main feature of the Gaussian distribution (the so-called bell-shaped distribution) can be entirely characterized by its mean and variance (Greene, 2002). A Paretian distribution does not show a well-behaved mean or variance. A power law, therefore, has no *average* that can be assumed to represent the typical features of the distribution and no finite standard deviations upon which to base confidence intervals (Moss, 2002). There are two major implications stemming from this.

One is that the dream of social science of building robust frameworks that allow social scientists to predict the evolution of social phenomena get shattered by the absence of statistical regularities in phenomena dominated by persistent interconnectivity. In fact, if stable mean and finite variance are absent, the probabilistic assessment of individual outcomes becomes much more difficult. This point reflects the more pervasive and structural issue of nonlinearity and emergence in complex systems. Linearity assumes the divisibility of systems into modules whose dynamics can be studied irrespective of the context. This point gives rise to the widely held independence assumption about phenomena that give rise to bell-shaped distributions.

The second is that power law tails decay more slowly than those of normal distributions. These ‘fat’ tails affect system behaviors in significant ways. For instance Buchanan (2004) reports that financial market drops of 10 percent in one day should occur once every 500 years according to a normal distribution. Mandelbrot (Mandelbrot and Hudson, 2004) shows that, instead, financial crises occur around once every five years. The lesson we can draw from this point is that extreme events, which in a Gaussian world could be safely ignored, are not only more common than expected but also of vastly larger magnitude and far more consequential.

Why Gaussian statistics mislead

A non-Gaussian world demands statistical methods that take into account path-dependency, nonlinearities, emergent properties of systems and the dynamics of multiple punctuated equilibria. The assumption of independence of events, which underlies the Gaussian world and the classical reductionist ‘variance theory’ approach (Mohr, 1982) leads to the wrong analytical tools and conclusions when dealing with connectionist dynamics (Kauffman, 1993; Holland, 1995). This is why, long ago, Mohr shifted to ‘process theory’ (see also Poole et al., 2000). This questioning of simple, variance-based reductionism leads one to also question the predictability and usefulness of studies based on independent events.

Greene’s textbook, *Econometric Analysis*, is in its 5th edition (2002) and is the standard for many econometricians and other social science researchers. He begins his ~950 pages of analysis with linear multiple regression and its five endemic assumptions:

1 independence among data points;
2 linear relationships among variables;
3 exogenous independent variables;
4 homoscedasticity and nonautocorrelation; and
5 normal distribution.

Mostly, the book focuses on how to make econometric methods work when one or more of these assumptions are untrue of the data. Given nonlinearity, for example, Greene says, ‘by using logarithms, exponentials, reciprocals, transcendental functions, polynomials, products, ratios, and so on, this “linear” model can be tailored to any number of situations' (p. 122). As for the normal distribution assumption:

> large sample results suggest that although the usual t and F statistics are still usable … they are viewed as approximations whose quality improves as the sample size increases … As n increases, the distribution … converges exactly to a normal distribution … This result is based on the central limit theorem and does not require normally distributed disturbances. (p. 105)

Greene observes that 'heteroscedasticity poses potentially severe problems for inferences based on least squares [regression analysis] ... It is useful to be able to test for homoscedasticity and if necessary, modify our estimation procedures accordingly' (p. 222). He then takes some 25 pages to discuss typically used methods to minimize the effect of varying variances: White test, Goldfeld-Quandt test, Breusch-Pagan/Godfrey LM Test, weighted least squares, two-step estimation, maximum likelihood estimation, model-based tests (i.e. analysis of residuals, Wald test, likelihood ratio test, Lagrange multiplier test, multiplicative and groupwise heteroscedasticity models), ARCH (autoregressive, conditionally heteroscedastic, Engle, 1982) model (three variants), with the generalized form, GARCH (Bollerslev, 1986), being most preferred. GARCH 'allows the variance to evolve over time' (p. 242). ARCH/GARCH assume that model errors appear in clusters and that the 'forecast error depends on the size of the previous disturbance' (p. 238) – it treats variance as a 'moving average of squared returns' (Engle, 1982).

Econometrics always assumes that data points are independent – always. Conditions calling for GARCH occur, but adjustments are made in modeling without ever giving up on the independence assumption. A plot of the GARCH moving average shows that any power law driven peak is whittled down to somewhat more than an average blip by the moving average process (see for example Ghysels et al., forthcoming, Figure 2). Ghysels et al. show that in ~90% percent of the extremes, even the best we have, GARCH, does not produce averages that account for extremes, and it clearly shows that no study based on averages can predict or account for extremes. MIDAS, an improvement on GARCH, which is newly defined in the Ghysels et al. article, is better at dealing with isolated extremes. More fundamentally, robustness compensations such as GARCH allow researchers to retain inappropriately narrowed standard deviations (by ignoring variance produced by the extremes), with the result that they risk claiming a level of statistical significance not validly descriptive of their data.

Greene does not discuss the Pareto, Zipf, Cauchy or Levy distributions. Nor does he discuss interdependent, interacting, connectionist, interconnecting, coevolutionary, or mutual causal data points, events or agents. Nor does he discuss when independence shifts to interdependence, or the reverse. These possibilities just do not exist in econometricians' assumptions about data. And yet theories underlying every kind of power law discovery all include a reference to interconnection of some form: power law phenomena invariably depend on *inter*dependent agents that, with some probability, are set off in a cycle of positive feedback or mutual causal progression resulting in an extreme event. In fact, none of the robustness adjustments to failing linear multiple regression assumptions that Greene discusses deal with the real world's *probable* – not just possible – losses of independence.

The bottom line is that the various robustness tests Greene discusses, even including the best and most widely used one, GARCH, give no assurance whatsoever that modern-day researchers account for the effects of extreme events in their statistical analyses. Let us put this in California earthquake terms, where we have insignificant quakes every day and a big one (where the ground moves 30 ft north) once every 150–200 years, with 6- and 7-level quakes occurring every decade or so. In effect, Greene and virtually all modern regression modelers of all stripes want Californians building and living in high-rise buildings to think that using a moving average (GARCH) of quake variance over the hundreds of harmless (average) quakes will lead to building codes that protect against the 8- and 9-level quakes. Anyone living in California and living through a significant quake will tell you this is nonsense. No amount of so-called robustness improvements to the standard linear multiple regression model allow it to model the effects of extreme quakes on buildings, bridges, lives and damage costs, that is, the effects of fat-tailed Pareto distributions. Needless to say, GARCH also does not accommodate the power law extremes that Mandelbrot has been observing in financial markets over the past 70 years (Mandelbrot and Hudson, 2004). *Robustness tests and solutions do not, and cannot, shift statistics from Gaussian to Paretian worlds.*

Consequences

We will not go through the entire list, but many organizational scholars have pointed to the growing disjunction between the Gaussian-based science appearing in journals and practitioner-oriented writing (Pfeffer, 1993; Anderson et al., 2001; Beer, 2001; Rynes et al., 2001; McKelvey, 2003; Van de Ven and Johnson, 2004). We are now in a position to describe the difference between attending to Gaussian and Paretian distributions. Virtually all of the statistics-based journal research rests on assumptions of independent events and Gaussian distributions. If one scans business-media books, such as *Organization and Environment* (Lawrence and Lorsch, 1967), *In Search of Excellence* (Peters and Waterman, 1982), *Images of Organization* (Morgan, 1986), *Rejuvenating the Mature Business* (Baden Fuller and Stopford, 1994), *Built to Last* (Collins and Porras, 1994), *Hidden Value* (O'Reilly and Pfeffer, 2000), *Good to Great* (Collins, 2001), *Knowledge Emergence* (Nonaka and Nishiguchi, 2001) and on and on, one sees that most of the cases and stories are about extreme events – successes or failures – but seldom about averages. No wonder there is a disjunction: managers live in the world of extremes; researchers using statistics report findings about averages. Talk about irrelevant research!

It is easy for people with no personal experience with an extreme event to think studies of averages are acceptable substitutes. People who go through tsunamis like the Christmas 2004 one off the Sumatran coast, earthquakes in California or Japan, hurricanes in Florida, floods along the Nile, Yangtze, Danube or Mississippi, or survive an avalanche in the Alps think differently, if they are still alive. Natural extremes are mostly negative. Organization extremes are both positive and negative. Early employees at Microsoft have one view of an extreme; those who were at Enron see theirs differently. The first thing we scholars have to do is get over the idea that studying averages is good science, is the only thing relevant to organizations and offers something useful to managers. Sometimes 'Yes', but we think 'Mostly No'.

The second task, once we face up to it, is to draw on more relevant disciplines, such as earthquake science where the study of extremes is routine, and complexity science, where focus is on emergent self-organization stemming from agent interdependence and positive feedback, consequent extremes and underlying scale-free theory (Andriani and McKelvey, 2005). We see very little in existing social science disciplines that offers anything constructive here. Only by facing up to this redirection of strategic organization research can it actually become a practioner-relevant science like the natural sciences. One of the lessons from earthquake science is that instead of studying all earthquakes, they study samples of 7s, 8s or 9s on the Richter scale. In point of fact, we have a large collection of case studies that are studies of extremes: those mentioned in the business-media books above and also in all of the MBA teaching cases. We even have multiple studies of single extremes, like a few 9s, that is, Xerox, the IBM pc, INTEL, ENRON and Parmalat, etc. With narrowed samples of similar extremes, Gaussian statistics and nonparametric methods are highly appropriate. Starbuck (n.d.) suggests a variety of other ways of 'learning from extreme cases'.

Our review of the different kinds of power law phenomena shows that underlying each is a collapse of the independence assumption. Once independence among data points collapses, and interdependence or interaction occurs, then the seeds of power law formations are planted. It is just a matter of time, just a matter of probability, for interdependent events to progress – from positive feedback effects – into an extreme event. As long as researchers look at the real world through the statistics lens, which means they have to make the independence assumption, the result will be Gaussian science and with it a denial of extreme events, a denial of infinite variance, a denial of unstable means – adding up to a denial of Paretian distributions. All of the denials simply narrow confidence intervals and wrongly allow researchers to claim statistical significance and then to assert their truth claim.

There are numerous conditions where natural data points do remain independent. For instance atoms and most molecules do not study, relate to, look at or learn from other atoms or molecules. In some cases, however, the imposition of energy past some critical point, for example Bénard's (1901) first critical value and resulting phase transition, turns even independent natural science data points into interdependent ones. In natural science, perhaps, scientists should still start with the null condition of independent data points. But in social science, where people do look at each other, do talk to each other, do learn from each other, do influence each other, etc., it seems to us that the null condition is one of interdependence.

Researchers should start with this assumption. They should start with the idea in mind that extreme events are a natural part of the social world.

No statistical findings, therefore, should be accepted if they gain significance by some assumption device by which extreme events and infinite variance are left out of the mix.

References

Anderson, N., Herriot, P. and Hodgkinson, G. P. (2001) 'The Practitioner–Researcher Divide in Industrial Work and Organizational (IWO) Psychology: Where Are We Now, and Where Do We Go From Here?', *Journal of Occupational and Organizational Psychology* 74: 391–411.

Andriani, P. (2003) 'The Emergence of Self-Organisation in Social Systems: The Case of the Geographic Industrial Clusters'. Unpublished PhD dissertation, Durham Business School, University of Durham, Durham, UK.

Andriani, P. and McKelvey, B. (2005) 'Power Law Phenomena in Organizations: Redirecting Organization Studies'. Academy of Management Meeting, Honolulu, Hawaii (August).

Baden Fuller, C. and Stopford, J. M. (1994) *Rejuvenating the Mature Business: The Competitive Challenge*. Boston, MA: Harvard Business School Press.

Beer, M. (2001) 'Why Management Research Findings Are Unimplementable: An Action Science Perspective.' *Reflections* 2: 58–65.

Bénard, H. (1901) 'Les Tourbillons Cellulaires dans une Nappe Liquide Transportant de la Chaleur par Convection en Régime Permanent', *Annales de Chimie et de Physique* 23: 62–144.

Bollerslev, T. (1986) 'Generalized Autoregressive Conditional Heteroscedasticity', *Journal of Econometrics* 31: 307–27.

Brock, W. A. (2000) 'Some Santa Fe Scenery', in D. Colander (ed.) *The Complexity Vision and the Teaching of Economics,* pp. 29–49. Cheltenham, UK: Edward Elgar.

Buchanan, M. (2004) 'Power Laws and the New Science of Complexity Management', *Strategy + Business Enews* 34: 1–8.

Casti, J. L. (1994) *Complexification: Explaining a Paradoxical World through the Science of Surprise*. London: Abacus.

Collins, J. C. (2001) *Good to Great: Why Some Companies Make the Leap...and Others Don't*. New York: HarperCollins.

Collins, J. C. and Porras, J. I. (1994) *Built to Last: Successful Habits of Visionary Companies*. New York: Harper Business.

Cook, W., Ormerod, P. and Cooper, E. (2004) 'Scaling Behavior in the Number of Criminal Acts Committed by Individuals' *Journal of Statistical Mechanics: Theory and Experiment* (5 July). Accessible at: http://ej.iop.org/EJ/journal/JSTAT

Engle, R. F. (1982) 'Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation', *Econometrica* 50: 987–1008.

Ghysels, E., Santa-Clara, P. and Valkanov, R. (forthcoming) 'There is a Risk-Return Tradeoff After All', *Journal of Financial Economics*.

Greene, W. H. (2002) *Econometric Analysis* (5th edn.) Englewood Cliffs, NJ: Prentice-Hall.

Holland, J. H. (1995) *Hidden Order: How Adaptation Builds Complexity*. Reading, MA: Addison-Wesley.

Kauffman, S. A. (1993) *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford: Oxford University Press.

Kaye, B. (1993) *Chaos and Complexity*. New York: VCH.

Lawrence, P. R. and Lorsch, J. W. (1967) *Organization and Environment: Managing Differentiation and Integration.* Boston, MA: Harvard University Press.

McKelvey, B. (2003) 'From Fields to Science', in R. Westwood and S. Clegg (eds) *Point/Counterpoint: Central Debates in Organization Theory,* pp. 47–73. Oxford: Blackwell.

Mandelbrot, B. B. and Hudson, R. L. (2004) *The (Mis)Behavior of Markets: A Fractal View of Risk, Ruin and Reward*. New York: Basic Books.

Mirowski, P. (1989) *More Heat than Light*. Cambridge: Cambridge University Press.

Mohr, L. B. (1982) *Explaining Organizational Behavior: The Limits and Possibilities of Theory and Research*. San Francisco, CA: Jossey-Bass.

Morgan, G. (1986) *Images of Organization*. Newbury Park, CA: Sage.

Moss, S. (2002) 'Policy Analysis from First Principles', *Proceedings of the National Academy of Sciences*, 99: 7267–74.

Nicolis, G. and Prigogine, I. (1989) *Exploring Complexity: An Introduction*. New York: W. H. Freeman.

Nonaka, I. and Nishiguchi, T. (2001) *Knowledge Emergence: Social, Technical, and Evolutionary Dimensions of Knowledge Creation*. Oxford: Oxford University Press.

O'Reilly, C. A. and Pfeffer, J. (2000) *Hidden Value: How Great Companies Achieve Extraordinary Results with Ordinary People*. Boston, MA: Harvard Business School Press.

Peters, T. J. and Waterman, R. H. (1982) *In Search of Excellence*. New York: Harper and Row.

Pfeffer, J. (1993) 'Barriers to the Advancement of Organizational Science: Paradigm Development as a Dependent Variable', *Academy of Management Review* 18: 599–620.

Poole, M. S., Van de Ven, A. H., Dooley, K. and Holmes, M. E. (2000) *Organizational Change and Innovation Processes: Theory and Methods for Research*. New York: Oxford University Press.

Rynes, S. L., Bartunek, J. M. and Daft, R. L. (2001) 'Across the Great Divide: Knowledge Creation and Transfer between Practitioners and Academics', *Academy of Management Journal* 44: 340–55.

Starbuck, W. H. (n.d.) 'Learning from Extreme Cases'. Accessible at: http//pages.stern.nyu.edu/~wstarbuc/

Van de Ven, A. H. and Johnson, P. E. (2004) 'Knowledge for Theory and Practice'. Working paper. Minneapolis, MN: Carlson School of Management, University of Minnesota.

West, G. B., Brown, J. H. and Enquist, B. J. (1997) 'A General Model for the Origin of Allometric Scaling Laws in Biology', *Science* 276: 122–6.

Zipf, G. K. (1949) *Human Behavior and the Principle of Least Effort*. New York: Hafner.

Bill McKelvey is professor of Strategic Organizing and Complexity Science at the UCLA Anderson School of Management. Current writing focuses on philosophy of science, organization science, complexity science, agent-based computational modeling and complex leadership. His book, *Organizational Systematics* (1982), remains the definitive treatment of organizational taxonomy and evolutionary theory. He has directed over 170 field study teams on six-month projects concerned with strategic and organizational improvements to client firms. He helped to found UCLA's Center for Human Complex Systems & Computational Social Science. Recent publications include: 'Model-centered Organization Science

Epistemology' (in J. A. C. Baum, ed., *Companion to Organizations*, 2002); 'Postmodernism vs. Truth in Management Theory' (*Research in the Sociology of Organizations*, 2003); 'Situated Learning Theory' (Y. Yuan, 1st author) (*Nonlinear Dynamics, Psychology, and Life Sciences*, 2004); 'Toward a 0th Law of Thermodynamics: Order-Creation Complexity Dynamics from Physics & Biology to Bioeconomics' (*Journal of Bioeconomics*, 2004); and 'Toward a Complexity Science of Entrepreneurship' (*Journal of Business Venturing*, 2004). *Address*: The UCLA Anderson School of Management, 110 Westwood Plaza, Los Angeles, CA, 90095–1481, USA. [email: mckelvey@anderson.ucla.edu]

Pierpaolo Andriani lectures in Management of Innovation at Durham Business School at the University of Durham, UK. His current research and publications focus on complexity theory and networks. In particular, he is interested in various theoretical aspects of complexity theory, such as power laws and fractals, and on the applications of complexity theory to explain the emergent properties of organizational networks (including industrial clusters). He is chairman of the Durham Business School Alumni Association, director of the sessions of the Summer School 'e-Business and Complexity: New Management Practices', director of the Complexity Society and Subject Editor of the international journal *Emergence: Complexity and Organisation*. *Address*: Durham Business School, Mill Hill Lane, Durham, DH1 3LB, UK. [email: pierpaolo.andriani@durham.ac.uk]