Organization Science
Vol. 20, No. 6, November–December 2009, pp. 1053–1071
ISSN 1047-7039 | EISSN 1526-5455 | 09 | 2006 | 1053

informs®

DOI 10.1287/orsc.1090.0481

PERSPECTIVE

From Gaussian to Paretian Thinking: Causes and Implications of Power Laws in Organizations

Pierpaolo Andriani
Durham Business School, Durham University, Durham DH1 3LB, United Kingdom, and
eBMS-ISUFI, Universita' del Salento, Italy, pierpaolo.andriani@durham.ac.uk

Bill McKelvey
UCLA Anderson School of Management, University of California, Los Angeles, Los Angeles, California 90095,
mckelvey@anderson.ucla.edu

Although normal distributions and related current quantitative methods are still relevant for some organizational research, the growing ubiquity of power laws signifies that Pareto rank/frequency distributions, fractals, and underlying scale-free theories are increasingly pervasive and valid characterizations of organizational dynamics. When they apply, researchers ignoring power-law effects risk drawing false conclusions and promulgating useless advice to practitioners. This is because what is important to most managers are the extremes they face, not the averages. We show that power laws are pervasive in the organizational world and present 15 scale-free theories that apply to organizations. Next we discuss research implications embedded in Pareto rank/frequency distributions and draw statistical and methodological implications.

Key words: Pareto; power law; Gaussian statistics; oganization; average; extreme events; scale-free theory
History: Published online in *Articles in Advance* October 7, 2009.

Although *normal* (bell-shaped) distributions and related quantitative methods are still relevant for a significant portion of organizational research, the increasing discovery of power laws signifies that Pareto rank/frequency distributions, fractals, and underlying scale-free (SF) theories are pervasive and valid characterizations of nonlinear organizational dynamics. Where they validly apply, researchers ignoring Pareto distributions risk drawing false conclusions and promulgating useless advice to practitioners. This is because under many circumstances what is important to most managers are the extremes they face, not the averages. Given this, we raise the question: *How do we redirect organization science toward the study of Pareto distributions in ways that still fall within the bounds of an effective science*?

Power laws signify Pareto rank/frequency distributions having *long and fat tails*, potentially infinite variance, unstable means, and unstable confidence intervals. Pareto distributions are alien to most quantitative organizational researchers, who are trained via Gaussian statistics to go to great lengths to configure their data to fit the requirements of linear regression models and related statistical methods. For example, most of the discussion in econometrics textbooks, such as Greene (2002), aims to accomplish this. Gaussian distributions have vanishing tails, thereby allowing focus to dwell on limited variance and stable means. As a result, confidence intervals for statistical significance are clearly defined, stable, and squeezed in toward the mean, increasingly the likelihood of achieving statistical significance.

The implications for organization science, however, go beyond extreme events. Tools do not exist in a theoretical vacuum. The adoption of normal distribution statistics carries a heavy burden of assumptions. Reliance on linearity, randomness, and equilibrium influences how theories are built, how legitimacy is conferred, and how research questions are formulated. Abbott (2001) says that linear thinking, what he calls the *General Linear Model*, defines the philosophical and methodological assumptions upon which linear science is based. It affects (a) how units of analysis are conceptualized, selected, and operationalized; (b) how variables are selected; and (c) how the interactions among variables are described by quantitative/qualitative models. Given these premises, we introduce *scalability* and the study of SF theories to begin a reorientation of the organization science paradigm from linear toward a *Pareto based science* more relevant to nonlinear organizational phenomena.

Scalability results from what Mandelbrot (1982) calls *fractal geometry*. A cauliflower is an obvious example. Cut off a floret, cut a smaller floret from the first floret, then a cut piece off the second, and so on. Now set them in line on a table. Each subcomponent is smaller than the former; each has the same shape and structure. They are fractal because they all look and behave about the same way. Fractals are signified by power laws and rank/frequency distributions. Researchers find organization-related power laws in intrafirm decisions, consumer sales, salaries, size of firms, ecosystems, director interlocks, biotech networks, and industrial districts, for example. These are all rank/frequency distributions.

Responding to the state of scientific disciplines of many kinds, Gell-Mann emphasized the study of "*surface complexity arising out of deep simplicity*" at the founding of the Santa Fe Institute (1988, p. 3; his italics). In describing the Santa Fe vision, Brock says the study of complexity "... tries to understand the forces that underlie the patterns or scaling laws that develop" as newly ordered systems emerge (2000, p. 30). Many complex systems tend to be *self-similar* across levels. That is, the same dynamics drive order-creation behaviors at multiple levels (West et al. 1997). These processes are called *scaling laws* because they represent dynamics appearing similarly at many orders of magnitude (Zipf 1949). We present fifteen SF theories, arguing that most apply to organizations. Gell-Mann (2002) argues that in living systems, scalability and scaling laws are as important a means of scientific explanation as is reductionism and explanation via law-like equations.

We first use findings from 141 kinds of power laws from natural to social and organizational phenomena to suggest the pervasiveness and importance of power laws, which typically signify well-formed rank/frequency Pareto distributions stemming from scalable causes. Next we classify 15 SF theories about scalable causal dynamics that apply to organizations, discussing several in detail. Then we switch to research implications: How do theory and methods change if we focus on rank/frequency Pareto distributions rather than squeezing all organizational phenomena into normal distributions (or more broadly, distributions that rely on finite variance)—as is currently the practice? Finally, we discuss implications in terms of the basic predictor function, $y = f(x) + \varepsilon$. How does basic thinking about prediction, data, error terms, and statistics have to change? A conclusion follows.

Entering the Third Phase of Complexity Science

Background

Complexity science has emerged in three phases.

Energy: The first phase appeared in Europe, led by Nobel Laureate Ilya Prigogine (1955; Prigogine and Stengers 1997). He built on Henri Bénard's (1901) study of emergent structures in fluids. Because these serve to dissipate energy imposing on a system, he labeled them *dissipative structures*. This phase transition, which occurs at the so-called *first critical value* of imposed energy, defines what we may call *the edge of order*. Schieve and Allen (1981), Haken (1983, 2004), Nicolis and Prigogine (1989), and Mainzer (2007) continue this line of work.

Emergence: This phase was initiated by Nobel Laureates Anderson (1972) and Gell-Mann (1988, 2002) along with Holland (1988, 2002), Kauffman (1993), and Arthur (1994) at the Santa Fe Institute. It is mostly oriented toward biology and the social sciences—i.e., living systems (Gell-Mann 2002). Its focus is on heterogeneous agents interacting at what was early on called *the edge of chaos*; this occurs at the *second critical value* of imposed energy. In between the "edges" of order and chaos is the region of emergent complexity, what Kauffman terms the *melting* zone (1993, p. 109). Bak (1996) argued that to survive, organisms have to have a capability of staying within the melting zone, maintaining themselves in a state of *self-organized criticality,* i.e., adaptive efficacy. Holland (2002) defines emergent phenomena as *multi-level hierarchies*, *intra- and inter-level causal processes*, and *nonlinearities*. Nonlinearity incorporates two additional outcomes: the *butterfly effect*[1] and *scalability*. Stacey (1992), Goldstein (1994), and many others apply complexity science to organization studies (Maguire et al. 2006).

Scalability: Though beginning decades ago with Pareto (1897), Auerbach (1913), and Zipf (1949), the third phase—which includes *econophysics* (West and Deering 1995, Mantegna and Stanley 2000)—focuses on power-law phenomena (Newman 2005). Econophysics began with Benoit Mandelbrot's (1963a) focus on stock market crashes. Although crashes are negative extreme events, their showing the power-law signature indicates that the markets were free to go up or down without restraint. *Power laws* often appear as *telltales of self-organization*, *emergence-in-action*, and self-organizing economies (Krugman 1996).

If one plots a well-formed Pareto rank/frequency distribution with both x and y axes as log scales, a negatively-sloped straight line will appear; this is the inverse power-law signature. Power laws often take the form of rank/frequency expressions such as $F \sim N^{-\beta}$, where F is frequency, N is rank (the variable), and β, the exponent, is constant.[2] This is in contrast to "exponential" equations stated in terms of the natural log, *e*, where the exponent is the variable and *N* is constant. Power laws show potentially infinite variance and an unstable or nonexisting mean and are frequently "... indicative of correlated, cooperative phenomena between groups of interacting agents ... " (Cook et al. 2004)—but not always, as we will point out below. Andriani and McKelvey (2007a) present more than 80 kinds of power laws (with sources)—16 physical; 24 biological; 21 social; and 23 pertaining to economic, business, and organizational phenomena. We show an expanded list of 101 social and organizational power laws in Table 1.

Power laws indicate long-tailed Paretian rank/frequency rather than "normal" Gaussian distributions—see Figure 1. The difference lies in assumptions about the correlations among events. In a Gaussian distribution, data points are assumed *independent-additive* (hereinafter simply *independent*). These events generate normal distributions, which sit at the heart of modern statistics. When causal elements are *independent-multiplicative*, a lognormal distribution results, which turns into a Pareto distribution as the causal complexity increases (West and Deering 1995)—detailed below. When events are *interdependent*, *interactive*, or both, normality in distributions is *not* the norm. Instead, Pareto distributions dominate because positive feedback (and other) processes leading to extreme events occur more frequently than "normal" Gaussian-based statistics lead us to expect. Further, as tension imposed on the data points increases to the limit, they can shift from independent to interdependent (Boisot and McKelvey 2007).

Phase three brings a totally new look to organizational applications of complexity science: (1) power laws as indicators of effective emergence-in-action, (2) SF theories as explanations of the underlying SF causal dynamics, and (3) Holland's "levers" as managerial action tools to foster scalable dynamics.

From Reductionism to a New Regularity: *Scalability*

Brock (2000, p. 29) says,

> The study of complexity... is the study of how a very complicated set of equations can generate some very simple patterns for certain parameter values. Complexity considers whether these patterns have a property of universality about them. Here we will call these patterns scaling laws.

The increasing discovery of power laws brings scalability and SF theories to prominence (Newman 2005). Many complex systems—resulting from emergent dynamics—tend to be *self-similar* across levels. That is, the same process drives order-creation behaviors across multiple levels of an emergent system (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). Scalability occurs when the appearance of phenomena is independent of the scale used to measure them (inches, feet, yards, miles) or the same causal dynamic operates at multiples levels.

Gell-Mann (2002) defines "effective complexity" as *regularities* or "schema" found or judged to be useful. For him, they appear as equations, genotypes, laws and traditions, and business best practices. What is new is Gell-Mann's recognition of a new regularity. In doing so, he sets forth two regularities:

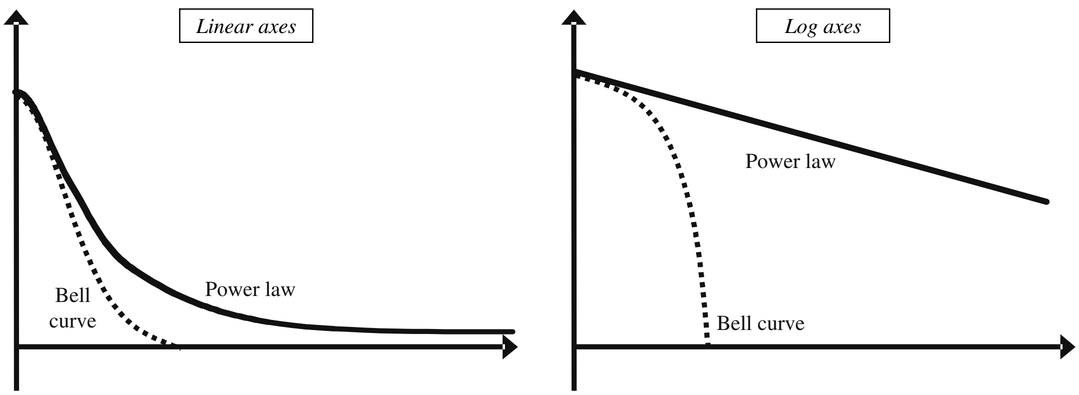
Type 1. *Reductionist Law-Like Regularities*: The *old simplicity* of reductionist causal processes of normal science, which are predictable and easily represented by equations—the data and information much preferred in classical physics and neoclassical economics (2002, p. 19). These are the point attractors of chaos theory—defined by forces, energy conservation, and equilibrium.

Type 2. *Multilevel SF Regularities*: The *new simplicity* of insignificant initiating events—what we call *butterfly events*. Outcomes over time that result from an accumulation of often random tiny initiating events that have lasting effects are compounded by positive-feedback effects over time, and become *frozen accidents* (2002, p. 20). These are the strange attractors and fractals of chaos theory—never repeating, fostering indeterminacy, and offering a different kind of regularity.

The first process generates regularities characterizing existing empirical organization and management research. These may be confidently described via Gaussian statistics and allow predictions that become the basis of schemata and prescriptive solutions. They are the basis of "reductionist" science—using components to explain a more macro level of behavior. The second focuses on the effects of tiny initiating butterfly events. The butterfly events of chaotic histories are seldom repeated, are not predictable, and can produce significant nonlinear outcomes that may become extreme events. Consequently, descriptions of these systems are at best problematic and easily outside the explanatory/scientific traditions of normal science. Gell-Mann concludes by noting that when butterfly events spiral up such that their effects appear at multiple levels and are magnified, we see self-similarity, scalability, and power laws.

Underlying most power laws is a causal dynamic explained via *SF theories*. Each theory points to a single pervasive generative cause to explain the dynamics at each of however many levels at which the scalability effect applies. SF theories yield what Gell-Mann (1988, p. 3) refers to as *deep simplicity*. Whereas tradition rests on the idea that lower-level dynamics can explain and predict higher-level phenomena and simplicity comes in the form of (usually) linear mathematical equations—i.e., reductionism (Gell-Mann 2002)—SF theories point to the same causes operating at multiple levels—the "simplicity" is one theory explaining dynamics at multiple levels. SF causes are Holland's *levers*. "... Almost all

Figure 1 Gaussian vs. Pareto Distributions



CAS [complex adaptive systems] exhibit *lever point* phenomena, where 'inexpensive' inputs cause major directed effects in the *C*AS dynamics" (2002, p. 29). These levers trigger *butterfly effects* across multiple levels.

Explaining via SF Theories

Along with those areas mentioned earlier, researchers find power laws in social networks, industry sectors, growth rates of firms, bankruptcies, transition economies, drug and movie profits, sales decays, and economic fluctuations—see Table 1. Power laws are mostly explained by SF theories. We identify 15 SF theories applying to organizations—see Table 2. We believe that the following logic chain applies:

1. Successful emergence results in fractals, SF dynamics, and power laws explainable via SF theory.
2. Power laws are far more ubiquitous than heretofore realized and are usually indicators of SF dynamics.
3. Consequently, SF dynamics are also ubiquitous; many SF theories seemingly apply to organizations.
4. If power laws are not obviously evident in organizations, then emergence has failed to emerge.
5. Therefore, organization-relevant complexity theory and research have to apply scalability dynamics.

Two new complexity thrusts are identifiable. First, roughly one-third of complexity science theory is missing in organizational and managerial applications to date, i.e., the scalability phase—power laws and the underlying fractals, scalability, and SF theory. Organizations are multilevel phenomena. Almost by definition we can take power-law signatures as the best evidence we have that emergence dynamics are operating at *multiple* organizational levels. We now know for sure that power laws apply at the overall industry level (Stanley et al. 1996, Axtell 2001) and industry sectors (Aoyama et al. 2009, Glaser 2009), with some appearing within firms. If power laws are not evident in a particular firm, we can only conclude that *emergence*, if it exists at all, is *not* multilevel. Building from the interacting food-web literature (Pimm 1982, Solé et al. 2001, Cuddington and Yodzis 2002, Sims et al. 2008), we can also surmise that absent the power-law signature, a firm's emergence dynamics are not capable of keeping it competitive with its changing competitors, suppliers, and customers (McKelvey et al. 2009). The bottom line is that power laws are significant indicators of crucially important managerial and organizational dynamics.

Second, organization change and entrepreneurship researchers should be especially interested in SF dynamics and related theories. Who more than entrepreneurs wouldn't like to let loose SF dynamics in their firms? Think of how many small entrepreneurial ventures stay that way simply because the emergent growth dynamics they had at the one- or two-level size failed to scale up as levels increased. Think how many large organizations show failing intrapreneurship for the same reason—the hundreds of "butterfly *ideas*" never become meaningful butterfly *events*, never produce butterfly *effects*, and never spiral into multilevel SF causal dynamics producing power-law signatures. Jean-Pierre Garnier, CEO of GlaxoSmithKline says:

> ...Size is a problem early in the drug-development process. "Drug finders" and innovators may well get tripped up by bureaucracy and tangled in red tape; good ideas are lost. Even worse, bad ideas may not be weeded out in time. (*The Economist* 2007, p. 57)

Complexity theory applied to organizations is silent on the foregoing points. One important move we recommend now is to learn how to apply SF complexity theories to organization change, organizational development, and entrepreneurship/intrapreneurship and strategy. Teaching and preaching complexity theory is out of date in our organizational world, absent SF theory. These points are further elaborated in Andriani and McKelvey (2007a, 2009) and Boisot and McKelvey (2007).

Causes of Power Laws in Organizations

Of the many kinds of power laws we list in Table 1, more than 50 are associated specifically with firms and organizational processes. Some of the power laws in

Table 1 Some Examples of Social and Organizational Power-Law Phenomena

1. Size of nations by population (Buldyrev et al. 2003)
2. Fractal structure of hunter/gatherer social networks (Hamilton et al. 2007).
3. Hierarchy of social group size (Zhou et al. 2005)
4. Economic fluctuations (Scheinkman and Woodford 1994)
5. Growth rate of countries' GDP (Lee et al. 1998)
6. Duration of recessions (Ormerod and Mounfield 2001)
7. Recessions and prosperity in Latin America (Redelico et al. 2008)
8. Transition economies (Podobnik et al. 2006)
9. Distribution of wealth (Pareto 1897, Levy and Solomon 1997)
10. Financial crashes (Sornette 2003)
11. Casualties in war (Cederman 2003)
12. Political complexity in communities (Carneiro 1987)
13. Size of cities (Zipf 1949)
14. Area, height, volume, size of buildings (Batty et al. 2008)
15. Costs of homeless in cities (Gladwell 2006)
16. Number of religious adherents (Clauset et al. 2007)
17. Price movements on exchanges (Mandelbrot and Hudson 2004)
18. Scientific discoveries (Plerou et al. 1999)
19. Copies of books sold (Hackett 1967)
20. Cascades in book sales (Sornette et al. 2004)
21. Sales of fast moving consumer goods (Moss 2002)
22. Movie profits (De Vany 2003)
23. Market share distribution of UK retail outlets (Moss 2002)
24. Cotton prices (Mandelbrot 1963a)
25. Blockbuster drugs (Buchanan 2004)
26. Wealth distribution of investors (Solomon and Richmond 2001)
27. Saving effects on wealth distribution (Patriarca et al. 2006)
28. Medieval wealth distribution (Hegyi et al. 2007)
29. Job vacancies (Gunz et al. 2001)
30. Changing language (Dahui et al. 2005)
31. Deaths of languages (Abrams and Strogatz 2003)
32. Social networks (Watts 2003)
33. Sexual networks (Liljeros et al. 2001)
34. Social influence (Castellano et al. 2000)
35. Coauthorships (Newman 2001)
36. Publications and citations (Lotka 1926, deSolla Price 1965)
37. Actor networks (Barabási and Bonabeau 2003)
38. Scale-free business networks (Souma et al. 2006)
39. Number of inventions in cities (Bettencourt et al. 2007)
40. Traffic jams (Nagel and Paczuski 1995)
41. Frequency of family names (Zanette and Manrubia 2001)
42. Global terrorism events (Dumé 2005)
43. Revenues of top 500 Chinese firms (Zhang et al. 2009)
44. Learning rates in heart surgery (Huesch 2009)
45. Firm size (Axtell 2001)
46. Firm size, interfirm relationships (Saito et al. 2007)
47. Growth rate: Japanese SIC industries (Ishikawa 2006)
48. Growth rates by sales: internal structure of firms (Stanley et al. 1996)
49. Growth rates: universities, countries (Stanley et al. 2000)
50. Economic effects of zero-rational agents (Ormerod et al. 2005)
51. Delinquency rates (Cook et al. 2004)
52. Aggressive behavior among boys (Warren et al. 2005)
53. Supply chains (Scheinkman and Woodford 1994)
54. Complex product development (Braha and Bar-Yam 2007)
55. Bankruptcy of firms (Fujiwara 2004)
56. Robustness in organizational networks (Dodds et al. 2003)
57. Learning strategy (Delaney et al. 1998)
58. Cognitive skills: "power law of practice" (Newell and Rosenbloom 1981)
59. Number of phone calls, emails (Aiello et al. 2000)
60. Website hits per day (Adamic and Huberman 2000)
61. News website visitation decay (Dezsõ et al. 2005)
62. "Fordist" power (Diatlov 2005)
63. Alliance networks among biotech firms (Barabási and Bonabeau 2003)
64. Branch networks of Polish firms (Chmiel et al. 2007)
65. Worldwide investment networks (Song et al. 2009)
66. Antibody alliances in biotech (Gay and Dousset 2005)
67. "Power curves" in U.S. industries (Zanini 2008)
68. Entrepreneurship/innovation (Poole et al. 2000)
69. Italian industrial districts (Andriani 2003)
70. Mergers and acquisitions waves (Park 2009)
71. Director interlock structure (Battiston and Catanzaro 2004)
72. Microsoft's ecosystem (Data from Iansiti and Levien 2004; additional analysis by Colon Drayton)
73. Market capitalization in industries (Glaser 2009)
74. Earnings, multilevel marketing by firms (Legara et al. 2008)
75. Biotech networks (Powell et al. 2005)
76. Growth of firms (Lee et al. 1998)
77. Productivity of innovation (Jones 2005)
78. Work incapacity from back pain (Schmid 2004)
79. Intra-firm decision events (Diatlov 2005)
80. Type of political officers, size of community (Johnson 1982)
81. Decision making and queuing (Barabási 2005)
82. Physical space, long-tail analysis (Bentley et al. 2008)
83. Japanese (J.) income[a]
84. J. income tax 1887–2003
85. J. firms' sales
86. J. firms' profit
87. J. company income
88. J. iron/steel fabrication sector
89. J. electrical machinery sector
90. J. wholesale sector
91. J. steel, other metals sector
92. J. general machinery sector
93. J. chemical, petroleum products sector
94. J. retail trade sector
95. France: size by total assets
96. France: size by sales in France
97. UK: size by total assets
98. UK: size by number of employees
99. Italy: size by total assets
100. Italy: size by sales in Italy
101. Italy: size by number of employees

[a]The source of the power-law distributions from 80–98 is H. Aoyama et al. (2009). Similar distributions have been found in many other industrialized countries—see for instance Gaffeo et al. (2003).

Table 2 Empirical Basis of Scale-Free Causes of Power Laws

Rules	Explanation
1. Surface-volume law	*Organisms; villages*: In organisms, surfaces absorbing energy grow by the square but the organism grows by the volume, resulting in an imbalance (Carneiro 1987); fractals emerge to bring surface/volume back into balance. West and Brown 2004 show that several phenomena in biology such as metabolic rate, height of trees, life span, etc., are described by an allometric power law whose exponent is a multiple of $\pm 1/4$. The cause is fractal distribution of resources. Allometric power laws hold across 27 orders of magnitude of mass.
2. Random walk	*Coin flipping; gambler's ruin*: Given a stochastic process such as coin flipping and, say, two players with a finite number of pennies to gamble, the probability that *eventually* one of the players will lose all his/her pennies is 100% (Kraitchik 1942). Number of tosses required is Pareto distributed (Newman 2005).
3. Hierarchical modularity	*Growth unit connectivity*: As cell fission occurs by the square, and connectivity increases by $[n(n-1)/2]$, producing an imbalance between the gains from fission versus the cost of maintaining connectivity; consequently, organisms form modules or cells so as to reduce the cost of connectivity; Simon argued that adaptive advantage goes to "nearly decomposable" systems (Simon 1962).
4. Event bursts	*Activity prioritization*: Individuals show bursts of communication, entertainment, and work activities followed by long delays, as opposed to random (Poisson) distribution (Barabási 2005).
5. Combination theory	*Number of exponentials; complexity*: Multiple exponential or lognormal distributions or increased complexity of components (subtasks, processes) sets up, which results in a power-law distribution (West and Deering 1995, Newman 2005).
6. Interactive breakage theory	*Wealth; mass extinctions/explosions*: A few independent elements having multiplicative effects produce lognormals; if the elements become interactive with positive-feedback loops materializing, a power law results; based on Kolmogorov's "breakage theory" of wealth creation (1941).
7. Interacting fractals	*Food web; firm and industry size*: The fractal structure of a species is based on the food web (S. Pimm quoted in Lewin 1992, p. 121), which is a function of the fractal structure of predators and niche resources (Preston 1948, Pimm 1982, Solé et al. 2001, West 2006).
8. Least effort	*Language; transition*: Word frequency is a function of ease of usage by both speaker/writer and listener/reader; this gives rise to Zipf's power law, now found to apply to language, firms, and economies in transition (Zipf 1949, Ishikawa 2006, Podobnik et al. 2006).
9. Preferential attachment	*Nodes; gravitational attraction*: Given newly arriving agents into a system, larger nodes with an enhanced propensity to attract agents will become disproportionately even larger, resulting in the power-law signature (Barabási 2002, Newman 2005).
10. Spontaneous order creation	*Heterogeneous agents* seeking out other agents to copy/learn from so as to improve fitness generate networks; there is some probability of positive feedback such that some networks become groups, and some groups form larger groups and hierarchies (Kauffman 1993, Holland 1995).
11. Irregularity generated gradients	*Coral growth; blockages*: Starting with a random, insignificant irregularity, coupled with positive feedback, the initial irregularity starts an autocatalytic process driven by emergent energy gradients, which results in the emergence of a niche. This explains the growth of coral reefs and innovation systems (Turner 2000, Odling-Smee et al. 2003).
12. Phase transition	*Turbulent flows*: Exogenous energy impositions cause autocatalytic, interaction effects and percolation transitions at a specific energy level—the first critical value—such that new interaction groupings form with a Pareto distribution (Prigogine 1955, Nicolis and Prigogine 1989).
13. Contagion bursts	*Epidemics; idea contagion*: Often, viruses are spread exponentially—each person coughs upon two others and the network expands geometrically. But changing *rates* of contagious flow of viruses, stories, and metaphors, because of changing *settings* such as almost empty or very crowded rooms and airplanes, result in bursts of contagion or spreading via increased interactions; these avalanches result in the power-law signature (Watts 2003, Baskin 2005) due to the small-world structures of the transient underlying networks.
14. Self-organized criticality	*Sandpiles; forests; heartbeats*: Under constant tension of some kind gravity (ecological balance, delivery of oxygen), some systems reach a critical state where they maintain adaptive stasis by preservative behaviors—such as sand avalanches, forest fires, changing heartbeat rate, species adaptation—which vary in size of effect according to a power law (Bak 1996).
15. Niche proliferation	*Markets*; when production, distribution, and search become cheap and easily available, markets develop a long tail of proliferating niches containing fewer and fewer customers; they become Paretian with mass market products at one end and a long tail of niches of the other (Anderson 2006).

Note. Additional power law causes are mentioned in West and Deering (1995), Sornette (2000), and Newman (2005).

broader social phenomena also apply to organizations. These discoveries of organization-relevant power laws offer substantial evidence that well-formed Pareto distributions are everyday organizational phenomena. Given this, two questions follow:

1. What causes power laws in organizations and what theories might explain the causes?
2. To what extent does the existence of power laws undermine prevailing assumptions that organizational phenomena are linear, equilibrium-seeking, and normally distributed?

In this section, we respond to the first question by arguing that SF theories apply to organizations.

Classifying Scale-Free Theories About Causes of Power Laws

We have assembled enough SF theories that a classification of them seems relevant, as follows:

1. *Ratio Imbalances—Theories* 1–4: In each, the basic SF cause is some kind of '*cost-driven efficiency* requiring constant or periodic adjustment:
2. *Multiple Distributions—Theories* 5–7: Here the SF cause of long tails is some kind of combination in the form of $p(y) \sim e^{a,b,c,d\cdots n}$, where the distributions underlying variables a, b, c, $d \cdots n$ are somewhat skewed and tainted with outliers. The multiplicative effect of the outliers progresses into a long-tailed Pareto distribution (West and Deering 1995).
3. *Positive Feedback—Theories* 8–11: In some systems the initial interaction possibilities are such that there is the possibility, if not probability, of positive-feedback spirals emerging simply as time progresses. The underlying SF cause is some probability that butterfly events will mutually interact so as to spiral up to produce long-tailed distributions.
4. *Contextual Effects—Theories* 12–15: Exogenous effects set SF dynamics in motion. In this set, different kinds of imposing effects set off SF causal processes. The common effect is context, but in each case the contextual effect is different and acts to set off a different kind of SF dynamic.

The definition and discipline base of each theory are given in Table 2. Ours is the first actual classification of this many theories about SF causes—most publications don't mention any SF cause; some mention one. Newman (2005) emphasizes preferential attachment and self-organized criticality, with minor reference to a few other physical ones (see also Sornette 2000). There are a few we don't include. Like the proliferation of power law discoveries, the growing set of SF theories makes it harder and harder not to wonder if they don't also apply to organizational phenomena.

Do Scale-Free Theories Apply to Organizations?

In Andriani and McKelvey (2007b) we make shorter arguments as to how most of the SF theories in Table 2 apply to organizations. Although each is briefly defined in Table 2, we discuss five of them here in more depth.

Square-cube/quarter-power. In biology, many scaling laws take the allometric form $Y \sim M^b$, where Y is some observable and M the mass of the organism. *Allometric* refers to a type of growth in which the parts of an organism grow at different rates determined by fixed ratios. Among these, West et al. (1997) cite metabolic rate, height of trees, life span, growth rate, heart rate, DNA nucleotide substitution rate, lengths of aortas, size of genomes, mass of cerebral gray matter, and density of mitochondria. In organisms, surfaces absorbing energy grow by the square but the organism grows by the volume, resulting in an imbalance. By adaptation, fractal structures emerge to keep surface absorption of energy in balance with the volume's use of energy. But energy is moved from surface to places in the volume by capillaries and other tubes in which fluid flows are governed by the *quarter-power* law. By nature, organisms adapt in a fashion such that quarter, square, and cube capacities are appropriately balanced. They are allometric'[3] scaling laws because they set up rigid relationships.

Allometric growth reflects universal structural constraints in the way organisms use energy. The emerging field of allometric growth (Whitfield 2006) is redefining biology and more in general the study about how ecosystems self-organize around fundamental energetic constraints. Insofar as organizational ecosystems use energy and energy-related quantities (money is the equivalent of energy according to some economists—see for instance Beinhocker 2007) and conform to general principles of ecosystem organization, we expect the study of allometric growth to yield compelling inputs to organization science.

Firms operate in ecosystems defined by the need to maximize revenues (exchange area between firm and customers) and minimize expenses (energy spent for developing, manufacturing, and distributing products). If this revenue-energy constraint can be given a meaningful geometric economic form, we may discover similar allometric relationships in many organizations. We now present some initial findings.

Haire (1959) first applied the square-cube law successfully to four firms. Levy and Donhowe (1962) confirmed his findings in 62 firms in eight industries. Stephan (1983) applied the square-cube law to firms in terms of effectiveness. Employees dealing with people outside the firm are "surface" employees—they bring in the resources from the environment. "Volume" employees are those inside who produce and coordinate. As firms grow, then, they have to maintain the square-cube ratio by adding more surface units or making them more efficient.

Carneiro (1987) applies the law to explain the upper bound on the size of villages. The law limits their size unless they develop what he terms "structural complexity," where complexity grows at two-thirds power of a village's population. Only by doing this do villages avoid splitting in two. Carneiro's theory is more general than Stephan's; Carneiro says social entities can increase in size only by building in structural complexity. In his data, for example, 100-person villages had 10 "complexity traits" whereas 1,000-person villages had four to five times as many. Johnson (1982), studying the governance of primitive organizations, finds that organizational complexity and leadership diversity scale accordingly with the allometric principles mentioned above. Much like Carneiro, Johnson finds that the emergence of nested hierarchical systems seems to be a response to the *scalar stress* induced by the exponential increase in the number of communication channels among the parts of the organization. The number of communication channels scale exponentially with the volume of the organization. A scalar stress increase forces the organization to elaborate more complex hierarchical systems with the effect of keeping scalar stress under control. By doing so, the organization changes the surface-volume ratio. From the point of view of the square-cube law, a decentralized network organization is a way to transform a large portion of the organizational employees into surface employees, thereby correcting scalar stress and bringing the surface/volume tension in line with allometric growth—but still subject to the quarter-power law.

The quarter-power law applies to the supply chain materials-flow "tube" that limits both the size of a supplier and retailer, for example. Zara is a retailer of new high-fashion designs—three weeks from models in the designer's mind to new fashions in its stores. How? By bringing manufacturing from China back to Europe, thereby shorting the "fluid" flow of clothing in transportation corridors and long-distance design communications. The primary factor in the virtual collapse of Citigroup stems from its "silo" design (Moore 2008). None of its employees having diverse vantage points of observations of its activities leading up to the subprime meltdown were connected to useful communication flow channels—no part of the firm could readily learn from any of the other parts.[4] One could call this "a total quarter-power breakdown."

Combination and Breakage Theories. Kolmogorov (1941) originally applied his "breakage theory" to coal—when large chunks of coal were smashed so as to be used in furnaces, resulting in small chunks down to powder—they appeared Pareto distributed. We now see Pareto distributions and power laws in chromosome breakage (Pevzner and Tesler 2003) and hydrodynamics (Bache 2004). In organizations the simple breaking up of firms into nearly autonomous modular designs makes breakage theory applicable—an approach dating back to Simon's (1962) "nearly decomposable" systems theory.

Oppositely, combination theory holds that the requirement for a power law to emerge is the number of elements in a complex system and their propensity to interact with one another. West and Deering (1995) and Newman (2005) both make the case that the combination of exponents results in a power-law distribution—the more of them that are combined, the more obvious the power law. Of course it does at the equation level, but what interacting non-normal phenomena actually occur in organizations?

If several organizational components or behaviors appearing vertically across several levels individually generate non-normal distributions having somewhat longer tails, and also influence each other, then combination theory tells us that organizations are inevitably going to contain well-formed Pareto distributions and show the power-law signature. Because they appear across several vertical levels, there is also a high probability that fractal structures and scalable causes are present—unless, of course, there are explicit attempts by management to negate them. In combination theory, the occurrence of interaction is taken as a naturally inherent likelihood as systems become more complex. Because many larger organizations have many degrees of freedom—and thus are complex by definition—they will show Pareto distributions, as is preliminarily evident in our Table 1.

As it has been applied to wealth (Montroll and Badger 1974), breakage theory appears as a set of independent-multiplicative elements that are lognormally distributed (West and Deering 1995, p. 152). To be wealthy, an individual has to have some *minimum* level of specified kinds of attributes (elements). The eight elements are *social background*, *education*, *personality type*, *technical ability*, *communication skills*, *motivation*, *right place-right time*, *willing to take risks*. But social background, education, technical ability, communication, and being in the right place at the right time are all potentially interactive, with an embedded positive-feedback effect—e.g., for a family, the more social status, education, and technical skills, then over time, the more technical skills, the more social status, etc. Now in the business world, suppose that some set of elements are required in organizations to cause the GEs, Microsofts, Toyotas, and Wal-Marts of this world—say *CEO skills*, *the right industry*, *new technology*, *"star" employees*, *special markets*, *weak competitors*, *borrowing ability*, etc. As these and other elements become more complexly interactive with positive-feedback effects, the distribution of firms changes from lognormal to Pareto. Axtell's (2001) research shows that this is, indeed, how U.S. industry appears.

Least effort. Zipf (1949) argued that *least effort* explained his Zipf's Law—a power law of word usage in English, French, and Spanish. The first question is, what is least-effort? Consider: Why would you learn words we don't use? Why would we use words you don't understand? In both cases it is wasted effort. Least-effort theory holds that each of us will minimize down to the only words relevant for meaningful transaction. For example, the 1953 Merriam-Webster's unabridged dictionary had more than 550,000 words; by 1971 it was reduced to about $450,000 words. Merriam-Webster's collegiate dictionary that most of us have in our offices is abridged down to around 86,000 words. The Harper-Collins Italian-English dictionary contains some 28,000 English words. We get by with fewer and fewer words.

Why is word usage Pareto distributed? Why are words like *the*, *of*, *and*, and so on at the top of the rank/frequency distribution? These words have to fit in with both the "before" words and the "after" words. Adjectives and adverbs, however, only have to fit with "after" words. Some words with narrow technical means are seldom used—the word "unabridged" is used only once in this paper until this sentence. High usage brought more opportunities, historically, for more proposed usages and more chances for disagreement on word usage. Higher usage also brings more opportunities for least-effort movements to improve efficiency. The result is increased demand for least effort and greater payoff. In word usage, then, we have an *interactive market transaction* of word usage that slowly works toward increased efficiency. The basic dynamic is a circular, positive-feedback process where each party moves toward the maximum efficiency, least-effort attractor basin. It is the opposite of the other positive-feedback SF theories.

To make least-effort theory even more compelling and applicable to organizations, we find that it has now shown to be especially characteristic of changing circumstances. Four recent studies suggest that Zipf's Law appears predominantly in the context of change:

1. In testing whether Zipf's Law applies to Chinese as well as English, Dahui et al. (2005) find that the power-law signature applies only during the period before Emperor Qin Shihuang's unification (about 1720), when Chinese characters were in flux. They conclude that the law does not apply when the number of characters is stable.

2. Ishikawa (2006) shows that Pareto's law holds when applied to firms in less populated JSIC[5] two-digit categories (having fewer firms) where growth rate is high, but a lognormal distribution applies to firms in large categories (filled with firms) where growth rate is, therefore, low.

3. Dahui et al. (2006) use two computational network models to show that the distribution of firms in growth markets is a power law but in markets without growth it is Gamma or exponential. They conclude, "...we cannot get [a] power-law distribution by preferential attachment in a constant market.... Economic growth is an important condition for the power-law distribution of firm size..." (pp. 363–364).

4. Podobnik et al. (2006) find empirically—and test further with a computational model—that time-series indices in transition economies (i.e., Hungary, Russia, Slovenia, etc.) fit Paretian rather than Gaussian distributions.

Three of the foregoing studies apply to organizations or markets.

Preferential attachment. This positive-feedback process (Barabási 2002) underlies biological and social networks, going from groups of individuals to groups of organizations. The Internet grows according to preferential attachment (Dorogovtsev and Mendes 2003). The same happens with cities and airport hubs—as nodes grow they attract even more people or flights (Barabási 2002). Internet marketing and sales are very much a positive-feedback process (Gladwell 2000). Any time a system grows by adding nodes to an existing network, the nodes' growth will amplify historically generated imbalances among the links. Absent top-down regulation, older or larger nodes will gain more links and generate a Pareto distribution—as in the biotech industry (Powell et al. 2005, Gay and Dousset 2005). Because organizations are made of social networks, preferential attachment plays a crucial role in their formation and evolution, thereby providing a solid base for a network-based theory of organizational formation and development. This "rich get richer" dynamic explains the emergence of central hubs and peripheral groups that characterize the geography of most organizations and the inherent concentration (and dispersion) of decision making. Other examples are Arthur's (1994) study of increasing returns—firms making profits can invest in things that make even more profits. Microsoft is a good modern example, as is Wal-Mart; the more it lowers prices, the more people come to buy; the more they come to buy, the more Wal-Mart can lower prices. And the more Wal-Mart sells, the more pressure it can put on suppliers to lower prices; the more they lower prices, the more Wal-Mart can sell.

Self-organized criticality (SOC). This theory is symbolized by Bak's (1996) and others' (Frigg 2003) sandpile experiments. A sandpile subjected to an infinitesimal external perturbation (sequentially adding single grains of irregularly edged sand) evolves toward a critical state, characterized by a *critical slope*, whereby any additional grain induces a systemic sand movement reaction that can span any order of magnitude (from one grain to thousands), with a frequency distribution expressed by a power law. This is counterintuitive. We generally assume a linear relationship between perturbation size and a system's reaction, i.e., small causes yield small effects. This is true before SOC is attained. Thus before criticality, each falling grain has a constant probability of displacing an adjacent grain. The probability of an avalanche therefore scales exponentially with the number of sand grains. This makes large avalanches highly unlikely. However, at *criticality* a power-law distribution results, given the global connectivity of the irregularly edged grains making up the sandpile. As Bak (1996, p. 60) writes, "In the critical state, the sandpile is the functional unit, not the grain of sand." SOC dynamics arise when an emergent system of links connects local pockets into a coevolving whole such that small and local fluctuations may be amplified to achieve systemic effects. More generally, as the tension in the system increases to the SOC limit (usually as a result of externally imposed tension—in Bak's SOC this is a function of gravity and accumulating sand grains) independent data points become interdependent. Mathematically this means that sandpile behavior obeys a power law of the type $F \sim S^{-\alpha}$, where F represents avalanche frequency with given size, S.

From the dynamics of earthquakes (Gutenberg and Richter 1944) and booms and busts in economic cycles (Sornette 2003, Mandelbrot and Hudson 2004, Malevergne and Sornette 2006) to the dynamics of supply chains (Scheinkman and Woodford 1994, Wycisk et al. 2008), a common pattern appears across disparate fields. Many systems exist in the state of criticality—on the critical slope, as it were. Bak argues that all systems in efficaciously adaptive states are in the state of criticality. Needless to say, then, SOC occurs frequently in markets and organizations (Buchanan 2000). Arguing that individual decisions are sticky like irregular sand grains, Bak applies SOC to economies. Because the tension between supply and demand builds and the actions to reduce it are not of equal size or regularity, free market economies operate at or near the critical state. Economic fluctuations (business cycles) are SOC (Scheinkman and Woodford 1994). We see SOC in the price of cotton and financial markets (Mandelbrot and Hudson 2004): many small changes in the price of stocks and the overall value of the market separated by volatility incidents averaging one in every four years from 1950 to 1980 and one every two years since 1980.[6] We also see SOC in consumer product sales (Moss 2002, Sornette et al. 2004) and managerial actions leading to different sized firms (Stanley et al. 1996, 2000)—all of which show power law signatures.

In the foregoing we detail how five SF theories apply to organizations. Elsewhere we argue that almost all apply (Andriani and McKelvey 2007b). Thus we have potentially 15 reasons why organization scientists should take as their ***new*** *null hypothesis* (Alderson 2008) the reality that organizations and managers very often live in a world of interdependent and not independent events; a world of Pareto distributions, fractals, and power laws; not Gaussian distributions where stable averages and finite variances across large samples are what count. Yes, we agree that there are many times and places where Gaussian statistics apply, but it is simply wrong to assume that they are the rule. If it is explicitly shown that a normal distribution holds, use Gaussian statistics. But absent this, the *new* null hypothesis should be presumed to apply.

Some Research Implications

> All the world believes it [Gaussian distribution] firmly, because the mathematicians imagine that it is a fact of observation and the observers that it is a theorem of mathematics. (Henry Poincaré 1913)[7]

We now offer Pareto driven alternatives, starting from a discussion of the predictor function. Take a standard predictor function consisting of a dependent variable, y, several independent or explanatory variables, x_n, and an error term, ε: Thus $y = f(x_1, x_2, x_3, \ldots, x_n) + \varepsilon$.

There are two concerns when one shifts from a Gaussian to a Paretian perception of data: (1) What happens to the predictor function? and (2) What happens to the error term? Organizational researchers using statistics as their basis of making truth claims—usually translated as findings significant at $p < 0.05$ or < 0.01—generally use statistical methods calling for Gaussian distributions.

For instance, Greene's textbook, *Econometric Analysis* (2002), is in its fifth edition and is the standard for many econometricians and other social science researchers. He begins his approximately 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 of error disturbances (p. 10). Mostly his 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). Regarding data distributions, he says (p. 105):

> 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.

Most standard econometric textbooks, such as Greene (2002) and Kennedy (2003), present methods to transform datasets into distributions with finite variance. Of these, the normal distribution is by far the most used due to its stability and conformance to the central limit theorem. However, as Bartels (1977, p. 86) writes:

> Economic data are seldom plentiful or accurate enough to distinguish between a hypothesized normal population

> and a nonnormal stable one, and since such data are notoriously long-tailed it is difficult to determine whether the population variance is finite or not.

This is crucial: Mandelbrot (1963b) claims that reliance on finite variance is the "Achilles heel" of econometrics. Our Table 1 offers reasonable evidence that increasing n may very well result in Pareto distributions—*the Central Limit Theorem* (*CLT*) *doesn't apply*! West (2007) takes it a step further, saying that Pareto distributions are so ubiquitous that finding a CLT-based average is "exceptional!"

The various robustness[8] tests standard econometric textbooks discuss give evidence that modern-day researchers have not taken on board Mandelbrot's (1963b, p. 438) plea:

> There is strong pragmatic reason to begin the study of economic distributions and time series by those that satisfy the law of Pareto. Since this category includes prices, firm sizes, and incomes, the study of Paretian laws is of fundamental importance in economic statistics.

Let us put this in California earthquake terms—about 16,000 insignificant quakes occur every year and a "really big one" once every 150–200 years, with 6- and 7-level quakes occurring within decades. If one sampled California quakes from 1995 to 2006, all but two would be in the 1–4 range: damage to no more than a few houses; no one killed. But this would miss the recent 6- and 7-level quakes in urban areas (costing billions of dollars and killing more than 100 people) and the next level 8 yet to come. Californians have long concluded that building codes should be based on the Pareto rather than Gaussian perspective. If California followed traditional econometric models, above, gathering more data could make quakes appear even more normally distributed, which is surely not the case.

In effect, application of methods based on Gaussian statistics (or more broadly on finite variance) models would lead Californians building and living in high-rise buildings to think that using a moving average of quake variance over the thousands of harmless (average) quakes would lead to effective building codes. Anyone living through a significant quake in California 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—i.e., the effects of fat-tailed Pareto distributions. *Robustness "solutions" cannot alter rank/frequency distributions to conform to Gaussian assumptions*.

The Predictor Function—From Gaussian to Pareto Thinking

Consider the typical "linear" prediction: $f(x_1, x_2, x_3, \ldots, x_n)$—the *predictor function*. Suppose we have a simple explanatory theory based on three independent variables: *Experienced, skilled, and satisfied employees increase output*. Thus $y = f(x_1, x_2, x_3) + \varepsilon$. In making a prediction like this, we usually *think* linearly—x causes y. Furthermore, propositions and even operational hypotheses appear in print with the expectation of a perfect correlation implicit—minus the effect of the error term. We visualize this as an upward sloping line in the wished-for plot of each cause of productivity, x_n (above), against output, y. Of course, the real world is never like this, and so the plot of y by x_1 data points, for example, appears as almost a circle at worst (near zero correlation) or a narrow ellipse at best—the thinner the better.

Two essential features are together the defining elements of a "normal" Gaussian distribution:

1. *The "mean" is stable and meaningful*: In the equation $y = f(x) + \varepsilon$, define y as weight and x as height. Average weight of males in the United States is 190 lbs.; average height is 5′8.6″. Millions of men are at or very near the mean.[9]

2. *Variance is finite*. Shortest living man is 2′5″; tallest is 7′10″ (both in China); both are within 1/2 magnitude.

In Gaussian statistics some variance is essential, but too much is a problem. Worse, if there is too much variance, confidence intervals widen and getting statistically significant results is more problematic. In our worker/output example, because human bodies are involved, independence is reasonable: There will be strong stable means of skill, experience, and satisfaction, with enough variance around the mean to allow correlations. But too little variance, and there is no meaningful correlation; too much variance and there is less (or no) chance for significance.

As one moves away from a simple study of bodies, such as our example—which is essentially where statistics-applied-to-firms started half a century ago—to study firms, distributions appear less obviously composed of independent data points. As a result, some 70 years of advances in statistics (since the founding of *Econometrica* in the early 1930s) offer devices econometricians can use to get all the weird kinds of data in the world of firms redesigned to fit linear regression. Now switch to a Paretian world. What changes? Consider species and consumer products.

- At one end of a Pareto distribution we have hundreds of elephants or Wal-Mart.
- At the other end we have trillions of mosquitoes or millions of "Ma & Pa" stores (defined as having no paid employees).

Elephants are huge but mosquitoes are tiny. Elephants eat vegetation, trample the land, and can trash your living room; mosquitoes suck warm blood, fly, and can give you viral diseases. Wal-Mart is huge,[10] enjoys substantial bargaining power over its supply chain,[11] and has powerful lobbying abilities.[12] Tiny Ma & Pa stores exist in microniches, buy supplies at the market, and have little if any political power. The averages of these *ends* of Pareto long tails are meaningless—Axtell (2008) says a "typical firm" doesn't exist! The mean, median, and mode are different. A careful Gaussian study at the average or median may offer little, if anything, of interest or use about firms at either end. A Gaussian study of Ma & Pa stores works but obscures microniches and offers little of value to Wal-Mart. And finally, any study close to the average or median (or elsewhere) ignores scalable dynamics.

What are the implications?

1. When extreme events occur they also alter the value of the mean—pulling it toward the tail where the extreme event occurred. Hence Pareto means are unstable.

2. Compared to "normal" variance, Paretian variance is potentially infinite. From Ma & Pa stores to Wal-Mart, profits, assets, and indebtedness range from thousands to billions of dollars. Profits and assets go from zero or worse to billions of dollars—crossing about 11 magnitudes. Mode, median, and mean are not the same. The larger the extremes, the less frequent or predictable they are. But when they happen, they increase the variance—perhaps more obviously in things like earthquakes and species abundance but also apparent in firms, merger and acquisition (M&A) activities, and bankruptcies.

3. Furthermore, because the variance is potentially infinite, the confidence intervals are considerably widened, making findings less apt to be significant. The CLT is meaningless.

Millions of businesses are single-proprietor or Ma & Pa stores with incomes in the thousands of dollars and possibly negative wealth. At the other extreme we see a few giant firms having hundreds of thousands of employees, billions of dollars in annual income, and assets of hundreds of billions—profits and assets can range across 11 magnitudes. For much organizational thinking, research, and practical applications, neither the distributions of the variables nor theorizing about the causal dynamics fits within the Gaussian assumption set. A different approach is needed, based on scalability.

Suppose we study 4,000 people in 500 small stores in small towns. Their owners' smartness; creativity; and knowledge of technology, markets, and customers, as well as good or bad attitudes, skills, behavior, networking, and so on, affect the other store owners and a couple hundred regular customers. This is not a bad sample, but if we improve it, both sample and error disturbances will become distributed more perfectly "normal." If we move the same study to 500 rural outlets of Wal-Mart we should end up with the same high quality "normal" distribution and error term disturbances. So far, so good.

But Wal-Mart is huge, having giant stores, many hierarchical levels, vast profits and assets, and global reach. So, we expand the sample to 1,000 worldwide. But now instead of two people at one level in each small store and 200 small-town customers to deal with, we have employees, acting at multiple levels in medium to giant stores, who have to deal with many more increasingly diverse customers, subordinates and superiors, local zoning issues, M&A issues, and so on. They make decisions ranging from local customer concerns to mid-management store policies affecting millions of dollars in profits to top-management policies and M&A activities with billions of dollars at stake. *No doubt, some aspects of human behavior in the N =1,000 remains normally distributed*. But as we include workers at each higher level of the hierarchy, things change. As we add levels, the dollar value of good and bad decisions increases: Some effects increase exponentially; some multiplicatively, and some may show interactivity and positive-feedback effects. Some of these skew distributions may combine to further assure Pareto-distribution effects.

What about timing? In small stores, decisions are pretty much the same from one year to the next. But at higher levels of Wal-Mart there are "routine" years at all levels and then some years where significant M&A, supplier realignment, or other decisions are made. We could sample across 5, even 10, years and miss the extreme outlier decisions such as buying the UK store chain that appears to be a mistake. If we study Wal-Mart people at the store-floor level, in one year, $N = 1,000$ will be normally distributed. If we study people at all levels across five years we *might* see a shift from normal to rank/frequency Pareto. But we may miss key extreme outliers; Wal-Mart doesn't make the really big decisions on a regular basis. To the extent our study includes people in larger and larger within-Wal-Mart networks and supply chain networks, involves multiple levels, and covers more years, all of the research issues embedded in rank/frequency research become more likely. It all depends on how much scalability is involved.

The Error Term

As noted above, once we plot Pareto distributed x and y on log scales, our expectation is a straight line. In empirical research, Greene (2002, pp. 7–8) observes that the clarity of a predicted relationship is clouded by the *normal distribution of error disturbances*—mostly, but not always, due to measurement error. These effects may be due to unknown or uncontrolled variables, measurement error, or both. How does this bear on Pareto based research?

Two points follow from the foregoing analysis. First, it is clear that in a Pareto rank/frequency world both the predictor function and error term are influenced by outliers that are fundamentally important to the validity of the analysis, as opposed to what are typically viewed as "throwaway" outliers in Gaussian statistics. In the latter, collecting a large sample almost inevitably means that the presumed validity of the analysis is improved—because the outliers have less effect, with statistical significance more easily obtained because of the narrower confidence intervals. In a Pareto world this is not true:

1. The analysis is faulty, if not totally meaningless, if the sampling of outliers is insufficient. In a Pareto world, building up the sample size while ignoring the outliers actually reduces the validity of the statistical analysis—it's like designing buildings based on average quakes while ignoring big ones.

2. Even though confidence intervals are widened, the power of the variance is in the long tails—meaning that if the outliers are properly sampled, the impact of the increased variance stemming from the tails more than compensates for the widening of the confidence intervals. For example, a Pareto distributed independent variable may be a strong predictor of a Pareto distributed dependent variable while leaving the error term *i.i.d.* (independent, identically distributed, as statisticians prefer); statistical significance is still relevant.

3. Correlations between Pareto and "normally" distributed phenomena are problematic; this needs further study.

Second, the concept of error rests on the "signal plus noise" paradigm introduced by Wiener (1949) as part of cybernetics. This long-standing paradigm is based on the assumptions that (1) a true measure of the signal exists as a deterministic function, (2) noise is random and its emergence is due to the system-environment coupling (and also because of measuring errors), and (3) the relationship between noise and signal is usually additive (Kennedy 2003, p. 8). Because noise is assumed to contain no relevant information about the system, filtering signal from noise is necessary to reconstruct the system's dynamical response.

In complex systems we have to rethink the signal-noise paradigm. The response of a complex system is a mix of order and disorder, represented in mathematical terms by deterministic and chaotic functions. Schroeder (1991) separates nonlinear "noise" into four colors, white (random), pink (deterministic chaotic), brown, and black (Paretian extremes).[13] In a complex system, chaotic fluctuations may reflect the fractal dimension of a system and its scaling properties. Consequently, "chaos" can be a fundamental part of the signal and may convey relevant information about its dynamics.[14] If this is true, then the basis for the distinction between signal (independent variables that are usually assumed deterministic and predictable) and noise (chaotic) becomes blurred (West 2006); consequently, the separability between signal and error term is called into question. In other words, if the signal is characterized by chaotic fluctuations that exhibit long-term correlations (as is usually the case for Paretian functions), and the separability between signal and noise cannot be based on the presence or absence of noise, then it is to be expected that the error term shows long-term correlations and "…long-term memory that ties events together" (West 2006, p. 271)—and it, therefore, is not Gaussian. This implies that statistical methods based on finite variance (i.e., classical regression models) may not be applicable when dealing with Paretian functions.

Some Methodological Implications

An SF theory approach in research starts from non-prejudicial views of the environment. Most conventional research depends on analytical functions and usually assumes linearity. Additional assumptions, again often implicit, concern evolutionary gradualism and equilibrium, with motion toward equilibrium considered adaptive in stable niches. Alternatively, research should start with a discussion about whether the phenomena under consideration show weak or strong interdependence among data points. If the former, then assume independence and the validity of calculus-based analytical functions. If the latter, then it is more likely the world is Paretian. Because fractals are continuous but infinitely irregular and therefore not amenable to differential calculus, the use of analytical functions becomes problematic (Mandelbrot and Hudson 2004). The little known mathematical fields of fractional calculus and Lévy-based statistics are more useful (West 2006). Thus new measures are needed. We focus on these next.

Develop Appropriate Measures of the Variables Relevant to SF Theories; Test for SF Dynamics:

(1) Start with our *new null hypothesis*. Determine whether a distribution is likely to be subjected to multiple dynamics, some of which may be Paretian, others Gaussian.[15] Two main questions here:

(a) Are the data points independent or interdependent? and

(b) Are the data points additive or multiplicative or interactive and scalable?

(2) If the answer is interdependent-multiplicative-interactive, then test whether interdependence increases going from small to large events. If yes, lognormal distributions likely could show Paretian tails. Then

(a) Don't exclude outliers. Even Pareto distributions may have inconsistent outliers (due to idiosyncratic causes); Sornette (2003) calls these "kings" or "black swans."

(b) Look for power-law signatures and identify the relevant parameters of the distributions.

(c) Is it a rank/frequency Pareto distribution?

(3) See if fractal structures exist:

(a) Study nestedness and self-similarity so as to establish fractal dimension(s);

(b) When looking at spatial, time independent phenomena (or time-dependent phenomena generated by distributed structures), look at the underlying generating network(s). Identify nodes and links and analyze the distribution of links to nodes. Calculate power-law slopes.

(c) Is the rank/frequency Pareto distribution well-formed, as indicated by a power-law distribution? It may signal emergence or not.

(d) For time series, determine fractal dimensions D: $D = 2 - \alpha$, where α is the scaling index of time series and α indicates the mix of ordered and random dynamics in the series (see Figure 2). $\alpha = 0.5$ indicates a completely random time series—there are no underlying patterns of order; it is purely random walk. $\alpha = 0$ or $\alpha = 1$ indicate a completely ordered phenomenon—no randomness anywhere. The interesting case occurs when $\alpha \neq 0$ or $\alpha \neq 1$; here the data show a mix of ordered and random dynamics that builds from previous fluctuations; the closer the value is to either end or to the middle, the more dominant the relevant dynamic is: order or randomness.[16] Emergence is most apt to occur as α goes below 0.25 or above 0.75. In Stanley et al. (1996) it is $\alpha = \sim 0.16$.

(e) For instance, Stanley et al. (1996) give an example of a hierarchical "Fordist" type organization where the CEO can order an increase in production, causing a Markov chain along the hierarchical levels—each subsequent action-step at time t is a replica of action at time $t - 1$. If it is carried out exactly from top to bottom of the firm, then the organization is strongly interdependent ($\alpha = 0$ for total top-down control, where α is the exponent of the power law describing growth variance), which means that the variance in growth rate is directly proportional to size. But lower-level managers and employees rarely follow orders exactly. If *all* ignore CEO orders, i.e., *all* act independently, then $\alpha = 1/2$. Usually employees follow orders with some probability and stickiness. Thus for a $0 < \alpha < 1/2$ or so (based on Stanley et al. 1996), we expect a power-law effect to obtain. Note that $0 < \alpha < 1/2$ could be due to a CEO's order implemented with some probability or it could be due to an emergent self-organizing process by employees.

Given Measures, then Consider the Following:

1. Develop theories and hypotheses based on SF theory that are aimed at causes or consequences of extreme events. The tools and measures mentioned above help identify the nature of the phenomenon and the appropriate SF theory (or mix of). See Table 2 for a list of SF theories.

2. Carry out empirical studies using data at frequently occurring scales—i.e., the hundreds or thousands of smaller events at lower-level scales comparable to the thousands of smaller quakes. Test whether these kinds of studies identify causes and consequences of larger extremes at, say, the next higher scale(s). That is, can we predict emerging fractal structures one level up in scale?

3. Because "extreme" extremes are rare in the real world, take a lesson from the econophysicists and use computational models to simulate known empirical findings and then test whether they stretch toward the more infrequent "extreme" extremes in the artificial computational world.

4. Work backwards from existing extreme events described in the organizational or managerial literature. We have already seen these sorts of studies carried out by official investigations of what led up to the Challenger and Pioneer disasters, the Bay of Pigs confrontation, Enron, 9/11, and so on. These findings then can be "reversed" and further tested by tracing backwards from extreme to smaller-scale employee networks and behaviors via computational modeling.

5. Use extreme-event statistics (Baum and McKelvey 2006) to calculate how extreme a future event might be. If a power-law tail is evident, one can do this simply by looking down the sloping line.

Conclusion

We suggest that fractals, rank/frequency Pareto distributions, power laws, and underlying scale-free theories will help organization scientists deal with Gell-Mann's "deep simplicity" (1988), scalability explanations of living systems in general (2002), and organizational complexity more specifically. We demonstrate that power laws are an inextricable aspect of how individuals, organizations, economies, and societies work. To answer the call for causal explanations relevant to organizations, we assemble a list of 15 scale-free theories and detail how several apply to organizations and management. These theories correct two key shortcomings of Gaussian research. First, they signify Pareto distributions and extreme events as elements of the managerial world that need to be accounted to by quantitative researchers; second, they put positive-feedback and other scale-free dynamics at the center of analysis.

Abbott's claim that the *General Linear Model* "subtly shaped sociologists"' thinking (2001, p. 7) (and the thinking of other disciplines such as economics, management theory, OB, etc.) may be at the base of the growing ineffectiveness between theory and practice. The gap between multiparadigmatic "science" appearing in journals and practitioner needs (Ghoshal 2005, Van de Ven and Johnson 2006) signifies the fact that the proliferation of academic disciplines has not produced research useful to practitioners (McKelvey 2006, McKelvey and Benbya 2007). Several environmental reasons may lie behind this reality: the ICT revolution, globalization, and radical transformations in Asia, for example, have contributed to the dazzling acceleration of change. These changes have increased global and local network connectivity making actors, from individuals to nation states, more interdependent and therefore more exposed to positive-feedback

Figure 2 Persistence and Anti-Persistence Behavior in Time Series

Order	Randomness	Order
0 ←	0.5	→ 1
Anti-persistence region (*like* followed with increasing probability by *unlike*) 0 = oscillatory pattern		Persistence region (*like* followed with increasing probability by *like*) 1 = ballistic trajectory)

dynamics and consequent rank/frequency distributions (Andriani and McKelvey 2007a).

Unfortunately, theories and methodological tools have not evolved at the same rate and are mostly still rooted in the time-honored concepts of equilibrium and linearity. In reality, organizational researchers study an interconnected world—full of rank/frequency discontinuities, chaotic dynamics, fractals, Pareto extremes, and power laws—with inappropriate research tools. The consequence is the gap between theory and practice that some theorists and many practitioners lament. In particular, theories and tools relying on "averages" and limited variance pledge allegiance to the altar of tradition—they force researchers to assume homogeneity instead of heterogeneity and averages instead of rank/frequency extremes.

The impact on use of statistics is significant. Researchers should start from the assumption that phenomena are rooted in interdependent dynamics and that long tails are the effect of scalable causal dynamics. Means and variance are unstable and cannot be used to represent the phenomenon, unless independence is demonstrated. We show that predictor and error terms acquire new meaning. We also show that complexity offers researchers some tools to characterize the mix of order and randomness in the systems, and we give examples about how research could be done in a Paretian world. More specifically,

1. Data about the trillions of mosquitoes or millions of Ma & Pa stores in one tail don't offer much useful information about the elephants or Wal-Mart in the opposite tail;
2. Methods that work on the large numbers in the Ma & Pa tail don't apply to studying extremes like Wal-Mart, Microsoft, Enron, or the organizational response behaviors to disasters like Katrina, Pioneer, and Challenger;
3. Which is to say, large samples at the mode don't speak to any other part of the distribution;
4. Studies of normal distributions at the median or mean don't speak to either tail;
5. Distributions of, and in, firms may not become "normal" just by increasing sample size;
6. Data collection working hard to include all Paretian outliers needs to replace approaches that delete outliers on the assumption that they are all errors and anomalies;
7. Scalable causes, dynamics, and theories become more important; they are absent from standard econometrics textbooks and current statistical practices in general;
8. Scalability-relevant methods simply don't exist in existing research approaches or in management theorizing.

The field of power-law science, extreme event theory, and complexity is relatively young. From the first Pareto distribution in Pareto's (1897) publication, Pareto rank/frequency and then power laws and scale-free theories have appeared in many instances. However, in comparison with the three centuries of development of the Newtonian/Gaussian world, power-law science is far from paradigmatic. There is no accepted standard for high quality research; limits of predictability are unknown; tools, frameworks, and methods are scarcely developed; the "line in the sand" that defines the spheres of influence of Gaussian and Paretian approaches needs clearer demarcation and new epistemological rules of justification logic.

Scale-free theories offer the promise of explaining extreme events and reducing the fragmenting effect of social science disciplines on organizational research. Discipline-centric researchers may dislike this consequence; discipline-neutral researchers will see research advantages and practitioner relevance. But remember: The *average* of the rank/frequencies from mosquitoes to elephants, from Ma & Pa to Wal-Mart retail firms, of from small aerospace-oriented foundries to Boeing and Airbus, or small computer repair stores to Microsoft, offers little useful information to any other part of a Pareto distribution. As Brunk says (2002, p. 36):

> Instead of the bulk of the data being produced by one process and the "outliers" by another, all events—both minuscule and the historically monumental—are produced by the *same* process in a SOC environment.

Whereas normal distributions call for more standardized management, the long unique tails of rank/frequency Pareto distributions call for more unique managerial responses. We argue that managers live in a world of mostly Paretian organizational and economic rank/frequency phenomena and that the fat/long tail and chaotic properties of Pareto distributions have to become more evident in empirical organizational research. For some portion of organizational research, the use of the so-called "robustness" enhancement techniques described in standard econometric textbooks is dysfunctional. Instead of being deleted, extreme events have to be properly sampled and analyzed. Given current quantitative practices, this is, indeed, a call for significant change. It is time to change.

Endnotes

[1]The so-called butterfly effect stems from Lorenz's (1972) paper: "Does the flap of a butterfly's wings in Brazil set off a tornado in Texas?" These are Holland's (2002) "tiny initiating events" that scale up to extreme outcomes.

[2]Though a power-law exponent is constant in a particular function, its exponent may change for different settings, industries, times, etc. Stanley et al. (1996) find slightly different scaling coefficients across a large sample of firms for sales, assets, number of employees, etc. Newman (2005) also shows different scaling coefficients.

[3]In general the exponent b is a multiple of $\pm 1/4$.

[4]"Besides ensuring that Citigroup has a proper handle on risk, Pandit's other challenge will be to streamline operations. Over the years, Citigroup has strapped together a vast array of businesses across its five business segments. The bank is now looking to improve efficiency and reduce overlap" (quoted from the *Morningstar* stock analyst report on Citigroup: http://news.morningstar.com/). In New York, for example, all the acquired businesses remained in their original, separate buildings—there was acquisition without integration, i.e., M&A without the "M."

[5]JSIC is the Japanese counterpart to the SIC code in the United States. Ishikawa (2006) studies all 14 Japanese two-digit industry classifications, which in the paper he refers to as "job categories."

[6]Just take a look at the market volatility chart in Ghysels et al. (2005) and count the number of times the red line (volatilities) goes above the black line, which represents the moving average (GARCH) line.

[7]Quoted in West and Deering (1995, p. 83).

[8]Other robustness techniques (not based on least square estimation) to deal with data sets that deviate from idealized assumptions can be found in Rousseeuw and Leroy (1987). In general, these techniques are not based on normal distribution and CLT but instead use the t-distribution. They assume finite variance, and like other robustness techniques, they have developed highly sophisticated tools (trimming, "winsorizing" to deal with outliers, skewness and long-tailed distributions tend to cut the tails start from trimming, winsorizing, etc.).

[9]From Wikipedia (http://en.wikipedia.org/wiki/Robust_Statistics accessed September 23, 2009).

[10]Wal-Mart is now the largest company in the world, has revenue more than $280 billion, and serves 138 million shoppers per year in approximately 5,300 stores (Bianco 2007).

[11]"Wal-Mart wields its power for just one purpose: to bring the lowest possible prices to its customers. At Wal-Mart, that goal is never reached. The retailer has a clear policy for suppliers: On basic products that don't change, the price Wal-Mart will pay, and will charge shoppers, must drop year after year. But what almost no one outside the world of Wal-Mart and its 21,000 suppliers knows is the high cost of those low prices. Wal-Mart has the power to squeeze profit-killing concessions from vendors. To survive in the face of its pricing demands, makers of everything from bras to bicycles to blue jeans have had to lay off employees and close U.S. plants in favor of outsourcing products from overseas" (Fishman 2003).

[12]Wal-Mart's lobbying expenses increased by 60% in 2007 (see Sarkar 2008).

[13]Following West (2006, p. 79) we define chaos as the "kind of randomness ... which is generated by the nonlinear dynamical property of a system." Chaos can be divided into deterministic chaos, colored and white noise, defined as follows: w*hite*, truly random, is characterized by a power spectrum whose exponent $\beta = 0$ (or frequency independent). Colored noise is divided into anti-persistent or mean-reverting [*pink*; deterministic chaos-based, anti-persistent; known as 1/*f* or power spectra with exponent $\sim f^{-1}$) and persistent (*brown* (f^{-2}) and *black* (persistent reoccurrence of extreme events; $f^{-\beta}$ with $\beta > 2$)] (Schroeder 1991). Colored noise and deterministic chaos can also be characterized by their dimensionality and pattern/path predictability (Dooley and Van de Ven 1999).

[14]The origin of chaos in complex systems' behavior is not always due to the system-environment coupling—although environmental interactions may contribute to it (Haken 1983)—but is often endogenous. Two consequences follow: First, chaos characterizes consistent dynamics of the system and therefore cannot be discarded as noise. Second, because chaos and noise are both nonlinear, separating them is problematic, though Dooley and Van de Ven (1999) and Baum and Silverman (2001) start down this path empirically. In Paretian systems a new type of mathematics and statistics is needed (West and Deering 1995, West 2006).

[15]Note that the Gaussian distribution belongs to a broader class of heavy-tailed distributions, the so-called Lévy stable distributions (West and Deering 1995). Lévy distributions need not be symmetric; they follow a generalized form of the law of large numbers. Lévy distributions are characterized by four parameters α, β, μ, c, where α (0–2) is the exponent, β represents the skewness, μ a scaling factor, and c a shift factor. For $0 < \alpha < 2$, we get the family of heavy-tailed distributions, which includes the Cauchy, Pareto, etc., most of which show no finite means and variance. $\alpha = 2$ yields the Gaussian distribution, which is a particular case of a much larger statistical distribution family. One referee of this paper notes that because the Gaussian distribution corresponds to a narrow region of the general class, decisions presuming the generality of Gaussian distributions risk being "brittle."

[16]More specifically: $0 < \alpha < 0.5$ indicates *anti-persistence* (see Figure 2). The system "remembers" a fluctuation and reacts with the opposite. Head is more likely to be followed by tail, a long stride by a shorter one, exploration by exploitation, or centralization by decentralization, etc.; $0.5 < \alpha < 1$ indicates *persistence*. In finance draw-downs and draw-ups (Sornette 2003) are repeated, i.e., sudden changes of stock market values that follow each other. For instance, the 1987 crash was really three financial crashes repeating (30.7% cumulative loss) in a short time period.

References

Abbott, A. 2001. *Time Matters: On Theory and Method*. University of Chicago Press, Chicago.

Abrams, D. M., S. H. Strogatz. 2003. Modelling the dynamics of language death. *Nature* **424** 900.

Adamic, L. A., B. A. Huberman. 2000. The nature of markets in the world wide web. *Quart. J. Electronic Commerce* **1** 5–12.

Aiello, W., F. Chung, L. Lu. 2000. A random graph model for massive graphs. *Proc. 32nd Annual ACM Sympos. Theory of Computing*, ACM, New York, 171–180.

Alderson, D. L. 2009. Catching the "network science" bug: Insight and opportunity for the operations researcher. *Oper. Res.* **56** 1047–1065.

Anderson, C. 2006. *The Long Tail*. Random House Business Books, London.

Anderson, P. W. 1972. More is different. *Science* **177** 393–396.

Andriani, P. 2003. The emergence of self-organisation in social systems: The case of the geographic industrial clusters. Unpublished doctoral dissertation, University of Durham, Durham, UK.

Andriani, P., B. McKelvey. 2007a. Beyond Gaussian averages: Redirecting organization science toward extreme events and power laws. *J. Internat. Bus. Studies* **38** 1212–1230.

Andriani, P., B. McKelvey. 2007b. Extremes and scale-free dynamics in organization science: Some theory, research, statistics, and power law implications. Presentation, Academy Management Conference, August 6, Philadelphia.

Andriani, P., B. McKelvey. 2009. Using Scale-free theory from complexity science to better manage risk. *Risk Management: An International Journal* **11**(4) Forthcoming.

Aoyama, H., Y. Fujiwara, W. Souma, Y. Ikeda, H. Iyetomi. 2009. *Pareto Corporations: Statistical Life and Death in Complex Business Networks*. Kyoto University, Kyoto, Japan. Forthcoming.

Arthur, W. B. 1994. *Increasing Returns and Path Dependence in the Economy*. University of Michigan Press, Ann Arbor.

Auerbach, F. 1913. Das Gesetz der Bevolkerungskoncentration. *Petermanns Geographische Mitteilungen* **59** 74–76.

Axtell, R. L. 2001. Zipf distribution of U.S. firm sizes. *Science* **293** 1818–1820.

Axtell, R. L. 2008. Nonexistence of a typical firm in the U.S. economy: Extremely heavy tails in firm size and growth. Presentation, Organizational Science Winter Conference, February 8, The Resort at Squaw Creek, CA.

Bache, D. H. 2004. From rupture to structure: The physical character of flocs. *Water Sci. Tech.* **50** 55–62.

Bak, P. 1996. *How Nature Works: The Science of Self-Organized Criticality*. Copernicus, New York.

Barabási, A.-L. 2002. *Linked: The New Science of Networks*. Perseus, Cambridge, MA.

Barabási, A.-L. 2005. The origin of bursts and heavy tails in human dynamics. *Nature* **435** 207–211.

Barabási, A.-L., E. Bonabeau. 2003. Scale-free networks. *Sci. Amer.* **288**(May) 60–69.

Bartels, R. 1977. On the use of limit theorem arguments in economic statistics. *Amer. Statistician* **31** 85–87.

Baskin, K. 2005. Complexity, stories, and knowing. *Emergence: Complexity and Organ.* **7** 32–40.

Battiston, S., M. Catanzaro. 2004. Statistical properties of corporate board and director networks. *Eur. Phys. J. B—Cond. Matt.* **38** 345–352.

Batty, M., R. Carvalho, A. Hudson-Smith, R. Milton, D. Smith, P. Steadman. 2008. Scaling and allometry in the building geometries of greater London. *Eur. Physical J. B* **63** 303–314.

Baum, J. A. C., B. McKelvey. 2006. Analysis of extremes in management studies. D. J. Ketchen, D. D. Bergh, eds., *Res. Method. in Strat. and Mgmt.*, Vol. 3. Elsevier, Oxford, UK, 123–197.

Baum, J. A. C., B. S. Silverman. 1999. Complexity in the dynamics of organizational founding and failure. M.R. Lissack, H.P. Gunz, eds. *Managing Complexity in Organizations*. Quorum Books, Westport, CT, 292–312.

Beinhocker, E. 2007. *The Origin of Wealth: Evolution, Complexity, and the Radical Remaking of Economics.* Harvard Business School Press, Cambridge, MA.

Bénard, H. 1901. Les tourbillons cellulaires dans une nappe liquide transportant de la chaleur par convection en régime permanent. *Ann. Chim. Phys*. **23** 62–144.

Bentley, R. A., M. E. Madsen, P. Ormerod. 2008. Physical space and long-tail markets. *Physica A* **388** 691–696.

Bettencourt, L. M. A., J. Lobo, D. Strumsky. 2007. Invention in the city: Increasing returns to scale in metropolitan patenting. *Res. Policy* **36** 107–120.

Bianco, A. 2007. *Wal-Mart: The Bully of Bentonville*. Doubleday Business, New York.

Boisot, M., B. McKelvey. 2007. Extreme events, power laws, and adaptation: Towards an econophysics of organization. *Best Paper Award. Acad. Mgmt. Conf.*, Philadelphia (August 7).

Braha, D., Y. Bar-Yam. 2007. The statistical mechanics of complex product development: Empirical and analytical results. *Management Sci.* **53** 1127–1145.

Brock, W. A. 2000. Some Santa Fe scenery. D. Colander, ed. *The Complexity Vision and the Teaching of Economics*. Edward Elgar, Cheltenham, UK, 29–49.

Brunk, G. G. 2002. Why are so many important events unpredictable? Self-organized criticality as the "Engine of History." *Japanese J. Poli. Sci.* **3** 25–44.

Buchanan, M. 2000. *Ubiquity: The Science of History… or Why the World Is Simpler Than We Think*. Weidenfeld and Nicolson, London.

Buchanan, M. 2004. Power laws and the new science of complexity management. *Strategy+Business* **34** 1–8.

Buldyrev, S. V, N. V. Dokholyan, S. Erramilli, M. Hong, J. Y. Kim, G. Malescio, H. E. Stanley. 2003. Hierarchy in social organization. *Physica A* **330** 653–659.

Carneiro, R. L. 1987. The evolution of complexity in human societies and its mathematical expression. *Internat. J. Comparative Soc.* **28** 111–128.

Castellano, C., M. Marsili, A. Vespignani. 2000. Nonequilibrium phase transition in a model for social influence. *Phys. Rev. Lett.* **85** 3536–3539.

Casti, J. L. 1994. *Complexification*. Abacus, London.

Cederman, L. E. 2003. Modeling the size of wars: From billiard balls to sandpiles. *Amer. Poli. Sci. Rev.* **97** 135–150.

Chmiel, A. M., J. Sienkiewicz, K. Suchecki, J. A. Holyst. 2007. Networks of companies and branches in Poland. *Physica A* **383** 134–138.

Clauset, A., C. R. Shalizi, M. E. J. Newman. 2007. Power-law distributions in empirical data. Accessed September 23, 2009, http://arxiv.org/abs/0706.1062v2.

Cook, W., P. Ormerod, E. Cooper. 2004. Scaling behaviour in the number of criminal acts committed by individuals. *J. Stat. Mech.: Theory and Exper.* http://www.iop.org/EJ/abstract/1742-5468/2004/07/P07003/JSTAT.

Cuddington, K., P. Yodzis. 2002. Predator-prey dynamics and movement in fractal environments. *Amer. Naturalist* **160** 119–134.

Dahui, W., Z. Li, D. Zengru. 2006. Bipartite produce-consumer networks and the size distribution of firms. *Phys. A* **363** 359–366.

Dahui, W., L. Menghui, D. Zengru. 2005. True reason for Zipf's law in language. *Phys. A* **358** 545–550.

Delaney, P. F., L. M. Reder, J. J. Staszewski, F. E. Ritter. 1998. The strategy-specific nature of improvement: The power law applies by strategy within task. *Psych. Sci.* **9** 1–7.

DeSolla Price, D. J. 1965. Networks of scientific papers. *Science* **149** 510–515.

De Vany, A. 2003. *Hollywood Economics*. Routledge, New York.

Dezső, Z., E. Almaas, A. Lukács, B. Rácz, I. Szakadát, A.-L. Barabási. 2005. Fifteen minutes of fame: The dynamics of information access on the web. http://mokk.bme.hu/archive/15fame_pr2005

Diatlov, V. 2005. Information technology in practice of organising: Meeting fragmentation and interdependence by the incremental political delivery of information systems in financial services. Unpublished Ph.D. dissertation, School of Management, University of Southampton, Southampton, UK.

Dodds, P. S., D. J. Watts, C. F. Sabel. 2003. Information exchange and the robustness of organizational networks. *Proc. Natl. Acad. Sci. USA* **100** 12516–12521.

Dooley, K. J., A. H. Van de Ven. 1999. Explaining complex organizational dynamics. *Organ. Sci.* **10** 358–372.

Dorogovtsev, S. N., J. F. F. Mendes. 2003. *Evolution of Networks*. Oxford University Press, Oxford, UK.

Dumé, B. 2005. Global terrorism follows a power law. *Physicsweb* February 1–2. http://physicsweb.org/articles/news/9/2/6/1.

Economist, The. 2007. The nimble sumo. (August 4) 57.

Fishman, C. 2003. The Wal-Mart you don't know. *FastCompany* (December), http://www.fastcompany.com/magazine/77/walmart.html.

Frigg, R. 2003. Self-organised criticality—What it is and what it isn't. *Stud. History Philos. Sci.* **34** 613–632.

Fujiwara, Y. 2004. Zipf law in firms bankruptcy. *Physica A* **337** 219–230.

Gaffeo, E., M. Gallegati, A. Palestrini. 2003. On the size distribution of firms: Additional evidence from the G7 countries. *Physica A* **324** 117–123.

Gay, B., B. Dousset. 2005. Innovation and network structural dynamics: Study of the alliance network of a major sector of the biotechnology industry. *Res. Policy* **34** 1457–1475.

Gell-Mann, M. 1988. The concept of the institute. D. Pines, ed. *Emerging Synthesis in Science*. Addison-Wesley, Boston, MA, 1–15.

Gell-Mann, M. 2002. What is complexity? A.Q. Curzio, M. Fortis, eds. *Complexity and Industrial Clusters.* Physica-Verlag, Heidelberg, Germany, 13–24.

Ghoshal, S. 2005. Bad management theories are destroying good management practices. *Acad. Management Learn Ed.* **4** 75–91.

Ghysels, E., P. Santa-Clara, R. Valkanov. 2005. There is a risk-return trade-off after all. *J. Financial Econom.* **76** 509–548.

Gladwell, M. 2000. *The Tipping Point: How Little Things Can Make a Big Difference.* Little, Brown, Boston.

Gladwell, M. 2006. Million-dollar Murray. *The New Yorker* (February 13), **82**(1) 96.

Glaser, P. 2009. Fitness and inequality in an increasing returns world. Working paper, UCLA Anderson School of Management, Los Angeles.

Goldstein, J. A. 1994. *The Unshackled Organization*. Productivity Press, Portland, OR.

Greene, W. H. 2002. *Econometric Analysis*, 5th ed. Prentice-Hall, Englewood Cliffs, NJ.

Gunz, H., B. B. Lichtenstein, R. Long. 2001. Self-organization in career systems: A view from complexity science. *M@n@gement* **5**(1) 63–88.

Gutenberg, B., R. F. Richter. 1944. Frequency of earthquakes in California. *Bull. Seismo. Soc. Amer.* **34** 185–188.

Hackett, A. P. 1967. *70 Years of Best Sellers*. Bowker, New York, 1895–1965.

Haire, M. 1959. Biological models and empirical histories of the growth of organizations. M. Haire, ed. *Modern Organization Theory*. Wiley, New York, 272–306.

Haken, H. 1983. *Synergetics, An Introduction*, 3rd ed. Springer-Verlag, Berlin.

Haken, H. 2004. *Synergetics: Introduction and Advanced Topics*. Springer-Verlag, Berlin.

Hamilton, M. J., B. T. Milne, R. S. Walker, O. Burger, J. H. Brown. 2007. The complex structure of hunter-gatherer social networks. *Proc. Royal Soc. London B* **274** 2195–2202.

Hegyi, G., Z. Néda, M. A. Santos. 2007. Wealth distribution and Pareto's law in the Hungarian medieval society. *Physica A* **380** 271–277.

Holland, J. H. 1988. The global economy as an adaptive system. P. W. Anderson, K. J. Arrow, D. Pines, eds. *The Economy as an Evolving Complex System*, Vol. 5. Addison-Wesley, Reading, MA, 117–124.

Holland, J. H. 1995. *Hidden Order*. Addison-Wesley, Reading, MA.

Holland, J. H. 2002. Complex adaptive systems and spontaneous emergence. A. Q. Curzio, M. Fortis, eds. *Complexity and Industrial Clusters.* Physica-Verlag, Heidelberg, Germany, 24–34.

Huesch, M. D. 2009. Power-law distributions in cardiac surgeon average volume? Working paper, Fuqua School of Business, Duke University, Durham, NC.

Iansiti, M., R. Levien. 2004. Strategy as ecology. *Harvard Bus. Rev.* **82**(3) 68–78.

Ishikawa, A. 2006. Pareto index induced from the scale of companies. *Physica A* **363** 367–376.

Johnson, G. A. 1982. Organizational structure and scalar stress. C. Renfrew, M. Rowlands, B. A. Segraves-Whallon, eds. *Theory and Explanation in Archaeology*. Academic Press, New York, 389–421.

Jones, C. 2005. The shape of production functions and the direction of technical change. *Quart. J. Economics* **120** 517–549.

Kauffman, S. A. 1993. *The Origins of Order*. Oxford University Press, Oxford, UK.

Kennedy, P. 2003. *A Guide to Econometrics*, 5th ed. MIT Press, Boston.

Kolmogorov, Á. N. 1941. Uher das logarithmisch Normale Verteilungsgesetz der Dimensionen der Teilchen hei Zerstuckelung. *Commits Rendus (DoLlady) de L'Academie des Sciences de l'URSS* **31** 99–101.

Kraitchik, M. 1942. *Mathematical Recreations*. Norton, New York.

Krugman, P. 1996. *The Self-Organizing Economy*. Blackwell, Malden, MA.

Lee, Y., L. A. N. Amaral, D. Canning, M. Meyer, H. E. Stanley. 1998. Universal features in the growth dynamics of complex organizations. *Physical Rev. Lett.* **81** 3275–3278.

Legara, E. F., C. Monterola, D. C. Juanico, M. Litong-Palima, C. Saloma. 2008. Earning potential in multilevel marketing enterprises. *Physica A* **387** 4889–4895.

Levy, S., G. Donhowe. 1962. Explanations of a biological model of industrial organization. *J. Bus.* **35** 335–342.

Levy, M., S. Solomon. 1997. New evidence for the power-law distribution of wealth. *Physica A* **242** 90–94.

Lewin, R. 1992. *Complexity: Life at the Edge of Chaos*. University of Chicago Press, Chicago.

Liljeros, F., C. R. Edling, L. A. N. Amaral, H. E. Stanley, Y. Aberg. 2001. The web of human sexual contacts. *Nature* **411** 907–908.

Lorenz, E. N. 1972. Predictability: Does the flap of a butterfly's wings in Brazil set off a tornado in Texas? Presentation, Meeting of the American Association for the Advancement of Science, Washington, DC.

Lotka, A. J. 1926. The frequency distribution of scientific productivity. *J. Washington Acad. Sci.* **16** 317–323.

Maguire, S., B. McKelvey, L. Mirabeau, N. Oztas. 2006. Complexity science and organizational studies. S. R. Clegg, C. Hardy, T. Lawrence, W. Nord, eds. *Handbook of Organizational Studies*, 2nd ed. Sage, Thousand Oaks, CA, 165–214.

Mainzer, K. 2007. *Thinking in Complexity*, 5th ed. Springer-Verlag, New York.

Malevergne, Y., D. Sornette 2006. *Extreme Financial Risks: From Dependence to Risk Management*. Springer-Verlag, Berlin.

Mandelbrot, B. B. 1963a. The variation of certain speculative prices. *J. Bus.* **36** 394–419.

Mandelbrot, B. B. 1963b. New methods in statistical economics. *J. Political Econom.* **71** 421–440.

Mandelbrot, B. B. 1982. *The Fractal Geometry of Nature*. Freeman, New York.

Mandelbrot, B. B, R. L. Hudson. 2004. *The (Mis)Behavior of Markets*. Profile, London.

Mantegna, R. N., H. E. Stanley. 2000. *An Introduction to Econophysics*. Cambridge University Press, Cambridge, UK.

McKelvey, B. 2006. Comment on Van De Ven and Johnson's "engaged scholarship": Nice try, but . . . . *Acad. Mgmt. Rev.* **31** 822–829.

McKelvey, B., H. Benbya. 2007. Using Pareto-based science to enhance knowledge for practical relevance. Presentation, Third Annual Organization Studies Summer Workshop, June 7–9, Rethymno, Greece.

McKelvey, B., B. B. Lichtenstein, P. Andriani. 2009. When organizations and ecosystems interact: Toward a law of requisite fractality in firms. M. J. L. Moreno, ed. *Chaos and Complexity in Organizations and Society.* UNESA, Madrid, Spain. Forthcoming.

Montroll, E. W., W. W. Badger. 1974. *Introduction to Quantitative Aspects of Social Phenomena*. Gordon and Breach, New York.

Moore, H. N. 2008. Vikram Pandit, Citigroup's chess master. *Online Financial News*, 8(May). http://www.efinancialnews.com/homepage/content/2450595389.

Moss, S. 2002. Policy analysis from first principles. *Proc. Natl. Acad. Sci. USA* **99** (Suppl. 3) 7267–7274.

Nagel, K., M. Paczuski. 1995. Emergent traffic jams. *Physics Rev. E* **51** 2909–2918.

Newell, A., P. S. Rosenbloom. 1981. Mechanisms of skill acquisition and the law of practice. J. R. Anderson, ed. *Cognitive Skills and Their Acquisition*. Erlbaum, Hillsdale, NJ, 1–55.

Newman, M. E. J. 2001. The structure of scientific collaboration networks. *Proc. Natl. Acad. Sci. USA* **98** 404–409.

Newman, M. E. J. 2005. Power laws, Pareto distributions and Zipf's law. *Contem. Phys.* **46** 323–351.

Nicolis, G., I. Prigogine. 1989. *Exploring Complexity: An Introduction*. Freeman, New York.

Odling-Smee, F. J., K. N. Laland, M. W. Feldman. 2003. *Niche Construction.* Princeton University Press, Princeton, NJ.

Ormerod, P., C. Mounfield. 2001. Power law distribution of the duration and magnitude of recessions in capitalist economies breakdown of scaling. *Physica A* **293** 573–582.

Ormerod, P., B. Rosewell, R. Colbaugh, K. Glass. 2005. Knowledge and cognition in the social sciences: Modeling with low cognition agents. Presentation, Conference on Human Complex Systems, May 19, Lake Arrowhead, CA.

Pareto, V. 1897. *Cours d'Economie Politique*. Rouge and Cie, Paris.

Park, J., B. Morel, R. Madnavan. 2009. Riding the wave: An integrative model of merger and acquisition patterns. Working paper. Katz Graduate School of Business, University of Pittsburgh, Pittsburgh.

Patriarca, M., A. Chakraborti, G. Germano. 2006. Influence of saving propensity on the power-law tail of the wealth distribution. *Physica A* **369** 723–736.

Pevzner, P., G. Tesler. 2003. Transforming men into mice. W. Miller, M. Vingron, S. Istrail, P. Pevzner, M. Waterman, eds. *Proceedings of the 7th Annual International Conference on Research in Computational Molecular Biology*. ACM Press, New York, 247–256.

Pimm, S. L. 1982. *Food Webs*. 2nd ed. University of Chicago Press, Chicago.

Plerou, B., L. S. N. Amaral, P. Gopikrishnan, M. Meyer, H. E. Stanley. 1999. Similarities between the growth dynamics of university research and of competitive economic activities. *Nature* **400** 433–437.

Podobnik, B., D. Fu, T. Jagric, I. Grosse, H. E. Stanley. 2006. Fractionally integrated process for transition economics. *Physica A* **362** 465–470.

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

Powell, W. W., D. R. White, K. W. Koput, J. Owen-Smith. 2005. Network dynamics and field evolution. *Amer. J. Socio.* **110** 1132–1205

Preston, F. W. 1948. The commonness, and rarity, of species. *Ecology* **29** 254–283.

Prigogine, I. 1955. *An Introduction to Thermodynamics of Irreversible Processes*. Thomas, Springfield, IL.

Prigogine, I., I. Stengers. 1997. *The End of Certainty*. Free Press, New York.

Redelico, F. O., A. N. Proto, M. Ausloos. 2008. Power law for the duration of recession and prosperity in Latin American countries. *Physica A* **397** 6330–6336.

Rousseeuw, P. J., A. M. Leroy. 1987. *Robust Regression and Outlier Detection.* Wiley, New York.

Saito, Y. U., T. Watanabe, M. Iwamura. 2007. Do larger firms have more interfirm relationships? *Physica A* **383** 158–163.

Sarkar, D. 2008. Wal-Mart lobbying up 60 percent in 2007. *Associated Press* (March 7).

Scheinkman, J., M. Woodford. 1994. Self-organized criticality and economic fluctuations. *American Econom. Rev.* **84** 417–421.

Schieve, W., P. M. Allen. 1981. *Self-Organization and Dissipative Structures.* University of Texas Press, Austin.

Schmid, L. 2004. Are living beings in a state of self-organized criticality? A new interpretation of data on work incapacity due to low back pain. *Medical Hypotheses* **62** 458–461.

Schroeder, M. 1991. *Fractals, Chaos, Power Laws*. Freeman, New York.

Simon, H. A. 1962. The architecture of complexity. *Proc. Amer. Philos. Soc.* **106** 467–482.

Sims, D. W., E. J. Southall, N. E. Humphries, G. C. Hays, C. J. A. Bradshaw, J. W. Pitchford, A. James, et al. 2008. Scaling laws of marine predator search behaviour. *Nature* **451** 1098–1103.

Solé, R. V., D. Alonso, J. Bascompte, S. C. Manrubia. 2001. On the fractal nature of ecological and macroevolutionary dynamics. *Fractals* **9** 1–16.

Solomon, S., P. Richmond. 2001. Power laws of wealth, market order volumes, and market returns. *Physica A* **299** 188–197.

Song, D.-M., Z.-Q. Jiang, W.-X. Zhou. 2009. Statistical properties of world investment networks. *Physica A* **388** 2450–2460.

Sornette, D. 2000. *Critical Phenomena in Natural Sciences*. Springer-Verlag, Heidelberg, Germany.

Sornette, D. 2003. *Why Stock Markets Crash?* Princeton University Press: Princeton, NJ.

Sornette, D., F. Deschâtres, T. Gilbert, Y. Ageon. 2004. Endogenous versus exogenous shocks in complex networks: An empirical test using book sale rankings. *Phys. Rev. Lett.* **93**(22). http://link.aps.org/doi/10.1103/PhysRevLett.93.228701.

Souma, W., H. Aoyama, Y. Fujiwara, Y. Ikeda, H. Iyetomi, T. Kaizoji. 2006. Correlation in business networks. *Physica A* **370** 151–155.

Stacey, R. D. 1992. *Managing Chaos*. Kogan Page, London.

Stanley, M. H. R., L. A. N. Amaral, P. Gopikrishnan, V. Plerou 2000. Scale invariance and universality of economic fluctuations. *Physica A* **283** 31–41.

Stanley, M. H. R., L. A. N. Amaral, S. V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M. A. Salinger, H. E. Stanley. 1996. Scaling behaviour in the growth of companies. *Nature* **379** 804–806.

Stephan, G. E. 1983. A research note on deriving the square-cube law of formal organizations from the theory of time-minimization. *Soc. Forces* **61** 847–854.

Turner, J. S. 2000. *The Extended Organism*. Harvard University Press, Cambridge, MA.

Van de Ven, A. H., P. E. Johnson. 2006. Knowledge for theory and practice. *Acad. Management Rev.* **31** 809–821.

Watts, D. 2003. *Six Degrees: The Science of a Connected Age*. Norton, New York.

Warren, K., G. Craciun, D. Anderson-Butcher. 2005. Everybody else is: Networks, power laws, and peer contagion in the aggressive recess behavior of elementary school boys. *Nonlinear Dynam. Psych. Life Sci.* **9** 155–173.

West, B. J. 2006. *Where Medicine Went Wrong.* World Scientific, Singapore.

West, B. J. 2007. The average person is truly exceptional: Where medicine went wrong. *17th Annual Conf. Society for Chaos Theory in Psych. and Life Sciences* (August 28). Chapman University, Anaheim, CA.

West, B. J, B. Deering. 1995. *The Lure of Modern Science: Fractal Thinking*. World Scientific, Singapore.

West, G. B., J. H. Brown. 2004. Life's universal scaling laws. *Phys. Today* **57**(9) 36–42.

West, G .B., J. H. Brown, B. J. Enquist. 1997. A general model for the origin of allometric scaling laws in biology. *Science* **276**(4) 122–126.

Wiener, N. 1949. *Time Series*. MIT Press, Cambridge, MA.

Whitfield, J. 2006. *In the Beat of a Heart: Life, Energy, and the Unity of Nature*. National Academy Press, New York.

Wycisk, C., B. McKelvey, M. Hülsmann. 2008. "Smart parts" logistics systems as complex adaptive systems. *Int. J. Phys. Distrib. Logist. Mgmt.* **38** 102–125.

Zanette, D. H., S. C. Manrubia. 2001. Vertical transmission of culture and the distribution of family names. *Physica A* **295** 1–8.

Zanini, M. 2008. Using "power curves" to assess industry dynamics. *McKinsey Quart.* (November) 1–6.

Zhang, J., Q. Chen, Y. Wang. 2009. Zipf distribution in top Chinese firms and an economic explanation. *Physica A* **388** 2020–2024.

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