Forecasting, When Power Law Distributions Apply

By

Richard John Logan

A thesis
submitted to the Victoria University of Wellington
in fulfilment of the requirements for the degree of
Master of Commerce and Administration,
Victoria University of Wellington

2012

Contents

Figures and Tables

Tables

Executive Summary

Forecasting, when power law distributions apply

Thesis focus:

The focus of my thesis, as outlined in my thesis proposal, is:

Whilst a lot of our strategic focus in the public sector is on linear policy approaches, many systems/ phenomena of importance are defined as non-linear or far from equilibrium. Traditional approaches to linear forecasting have not proved effective for non-linear systems, since non-linear systems follow a different set of rules. Historically, non-linear systems were too hard to forecast, but over recent decades some rules and approaches are starting to emerge.

One important and clearly defined category of non-linear systems are those that follow a 'power-law' distribution rather than the 'normal' distribution, which is often associated with linear systems or systems in equilibrium. My research collects, analyses, and does a comparative analysis of the different power law populations, as well as the main strategic forecasting techniques that can be applied to those populations/ systems.

Overall Conclusions and observations

Just as in science and mathematics, there is now a clearly defined separation and understanding of linear and non-linear systems and the rules that apply to each. My thesis has as its central theme, the idea that strategy as a subject also fits this same philosophical separation of approaches, which I have called the strategic planning versus the strategic thinking divide. Strategic planning is essentially the linear approach – being rational and assuming relatively stable conditions. Strategic thinking assumes the world is effectively non-linear and 'far from equilibrium'. Non-linear approaches mean acknowledging concepts like; punctuated equilibrium, power law 'log-log' graphs, 'scale-free' characteristics, 'self organising criticality', accepting only pattern prediction (including 1/f formulas) and not precise prediction etc. Understanding non-linearity is essential to understand such things as 'Black Swans'. Luck, serendipity and 'bounded rationality' are always involved in non-linear complex adaptive systems, whereas linear systems tend to comply with the so called 'rational' traditions in science and economics.

Power law statistical distributions can be seen in a wide variety of non-linear natural and man-made phenomena, from earthquakes and solar flares to populations of cities and sales of books. This sheer diversity of effects that have power law distributions is actually an amazing fact that has only become evident over the last decade or so. Since the world contains aspects that are clearly linear and other aspects that are clearly non-linear, it is essential for someone interested in strategy to be able to understand both systems and be able to apply

the correct techniques to each approach. The two parts of 'punctuated equilibrium' effectively link the two strategic approaches together as there is only one world and not two separate realities.

It therefore follows that a strategist needs a good understanding of both strategic planning and strategic thinking, since both are needed for different phases or periods, and perhaps both are needed for any period when you can't tell what phase you are in, which can also happen. I suggest that under a linear phase, the strategic planning approach should be dominant, but supported by strategic thinking (since you never know when events will turn abruptly); whereas in a turbulent non-linear period the strategic thinking approach should be dominant, but supported by strategic planning (since you know that great turbulence will not last). This is a sort of a swapping dominant/ recessive situation, which has a loose parallel in the theory of the left/ right brain split, where it is not wise to use only one style of thinking, since there are two styles which suit different situations. The key is to pick the right thinking style for the right situation. Just as we have one brain, but two thinking styles, so in the strategy toolbox we also have two valid, useful and complimentary general strategic approaches.

However for this thesis, I have focused on the non-linear power law aspects of life which have strong implications for strategic thinking, since that is the new area for me as well as one of the new knowledge frontiers for strategy as a subject (and for leadership, politics and many other areas).

Specific observations

- **The 'sandpile' metaphor is the key non-linear metaphor**
  The answer to my research question, 'Forecasting, when power law distributions apply', is to understand the 'sandpile' metaphor of Per Bak, where the next event could be very small or occasionally very large, and that the overall pattern of the events will follow an emergent power law 'log-log' chart. From this follow mathematical $1/f$ type ratios. Inherent in the sandpile metaphor is the associated concept of self-organising criticality, in which a scale-factor of a system diverges, because the system automatically drives itself to a critical point by some dynamic process, no matter what state they start off in. The divergence can leave the system with no appropriate scale factor to set the size of some measured quantity and therefore the quantity follows a power law. The power-law distribution is the only distribution that is scale-free.

- **Power laws distributions are inherently unfair (the 'rich-get-richer')**
  The Pareto process is a rich-get-richer mechanism in which the most populous cities or best-selling books get more inhabitants or sales in proportion to the numbers they already

8

have. These distributions have a power law in their tail. These 'rich-get-richer' processes are also known as the Matthew effect, cumulative advantage or preferential attachment. The **'Matthew' effect** comes from the biblical edict, *'For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath.'* (Matthew 25:29 – King James Bible). To me this edict seems harsh, but is very apt for reality. Newman believes that this principle is now the most widely accepted theory for explaining such things as citations, city populations and personal income.

- **Increased networks/ interdependence, means increased turbulence**
  As the world becomes more connected through social, political and economic networks, the importance of understanding non-linear systems increases as turbulence in all these systems dramatically increases due to power law effects within network flows. Extremes will get more extreme due to the 'Matthew' effect, where the winners take all.

- **Accepting that power laws only give pattern prediction not point prediction**
  For example, you will never be able to predict the size of the next earthquake or war, but the power law graph for that subject gives a clear forecasting framework at the highest level looking at all the available data, which we can learn from and help us get over the idea of needing to predict the next event. For example, Karl Popper believed that under a 'punctuated equilibrium' scenario, forecasting inherently is not possible, but you should at least understand why that is so.

- **No difference between natural and human power law activities**
  For this thesis (which is a preliminary study of the subject) I have found no effective difference between natural effects (e.g. earthquakes or animal extinction rates) that follow power law distributions and human activities (e.g. wealth, war, citations etc) that follow power law distributions. I have therefore not separated the two different types in this thesis. In fact the later section on networks is in purely alphabetical order of the subject. This lack of difference is not a comforting finding. The implications of this lack of clear difference between natural and human effects is considered in the Conclusions chapter. It is a profound question, as to whether we can actually change the various underlying power law patterns or are we predestined to fit these patterns.

- **Complexity is still an emerging subject, against a background of linear thinking**
  Like all emerging areas, the study of complexity (non-linearity) naturally comes up against resistance from the established paradigms of thinking which are essentially linear and equilibrium in focus. For the subjects that I covered, this reaction is most apparent in traditional (meaning much of current) economics. For most items covered I put forward both views - the traditional and the complexity explanations. Studying so many different

power law examples was important to reinforce how widespread and conceptually similar all these different power law examples all are, even if they are describing completely different things.

- **Sample sizes for power law graphs are impressive**
  For example, many of the populations tested are the complete population at a given point or for a period of time. Sometimes the sample sizes are measured in millions of data points. For example, Stanley collected 35 million observations on the Paris Stock Exchange. Since these are impressive sample sizes or are the complete populations for a given time period, it is clear that any thoughts of my sampling any population is by comparison slightly futile.

Importance of modelling

One very important tool, which I have hardly mentioned in this thesis, is 'modelling'. Probably every power-law example in this thesis comes from sources that not only record the power-law log-log chart but also record attempts at modelling that subject using that chart's power-law co-efficient. I acknowledge the extreme importance of power-law charts to modellers since this has provided new methods and a greatly enlarged range of subjects (i.e. everything non-linear) to model. As stated in the introduction, my interest is only in the high level principles of non-linearity as it affects strategy. I am not a modeller, so I have not emphasized that aspect but that should not be taken as a lack of appreciation of how important power-laws are to modelling over a whole range of subjects. Whilst 'modelling' is an important technique it is not the one word answer to my thesis question. It is not my interest or my journey for this thesis. However, I believe this thesis would be a good background explanation of the subject of power-law non-linearity for modellers and non-modellers (like me).   It doesn't help you 'predict' the next event, but it should provide a good description of what is happening or could happen.

Power law applications and key non-linear concepts

In this thesis, I have collected and analysed as many power law distributions (log-log charts) as I can find. Whilst the reader may find this repetitive, I found this repetition extremely important because it reinforced how widespread power-law effects are and because I applied the 'emergent' concept to the subject of non-linearity itself. For example, I knew quite early on that each example was often incomplete in itself, but collectively the reader could see that the same themes and the same linear versus non-linear debates were occurring again and again. Hence the whole (different examples of power-law effects) is in my view greater than the sum of the incomplete individual parts. I note that there are other non-linear distributions (which I have not studied) that are not power-law distributions and that Newman and others have indicated that not all distributions that are said to be power-law distributions are power-law.

Throughout the thesis the main ideas and concepts that strategic forecasting needs to apply to non-linear power-law distributions are:

- Understanding the importance of complex adaptive systems,
- 'Scale-free' or 'scale invariant' features,
- Punctuated equilibrium (periods of stability and periods of rapid change),
- Self-organised criticality (poised at a tipping point),
- The 'Butterfly' effect (small changes can have big effects),
- Emergent properties or patterns,
- The 'Matthew' effect (winner takes all),
- Change is caused by both Internal/ external triggers,
- Pattern prediction rather than point prediction,
- Straight line log-log graphs are the signature of power law-distributions,
- With emergent power-law patterns, details don't matter,
- Increased networks/ interdependence, means increased turbulence, and
- Strategy is about surviving the 'Red Queen' Race.

These concepts are covered throughout the thesis as they apply to each power law application and they are summarised in the Conclusions chapter.

Forecasting, when power law distributions apply

Chapter 1- Introduction

Background to my thesis

My research topic builds on aspects of my Master of Strategic Studies (MSS) course work, especially STRA 519 (Research Essay), STRA 511 (Strategic Thinking) and STRA 502 (Strategic Analysis).

In my 2009/10 research essay for STRA 519, 'Picking the best strategic forecasting tool for the job at hand', I mostly focussed my efforts on what I would now describe as 'linear' approaches that suggested one could forecast into the future (such as forecasting one or more scenarios). These approaches included: force field analysis, induction, deduction, extrapolation, theoretical and judgemental approaches. I also considered dimensions of change such as incremental, substantial and transformational change. I covered these areas further than I had before and in my mind was comfortable that I understood the different techniques that I had considered. I also touched on the effect of power laws, but since that was a new area of knowledge for me, I did not have time or space to explore that area further or in any depth. Understanding power laws and complexity was therefore an area of unfinished business as far as my learning was involved.

It is therefore natural that just as I finished my STRA 519 research essay, I realised that I had only finished half of it and I had what for me, was the unexplored area of non-linear forecasting to do and that was in my mind like a part two of my first essay. The STRA 519 research essay summarised approaches that believed the future can to some extent be forecast, whereas this essay will be the opposite, that forecasting the future is not possible on all important matters, especially when dealing with complex adaptive systems. This is because of concepts such as tipping points, punctuated equilibrium, emergent properties, butterfly effects and scale free effects. Whilst I had some familiarity with many of these terms; I was aware that these and other non-linear concepts (also called 'complexity'), use a completely separate set of jargon from the books that I had previously been studying. I have used this thesis to gain an understanding of this complexity language.

Throughout my Master of Strategic Studies degree I started to notice that many of the articles in our reading packs talked about the difference between thinking about linear things versus thinking about non-linear things. For example, in STRA 511 – Strategic Thinking, I looked long and hard at the readings provided, by thinkers like F.A. Hayek (complex systems) and John Sterman (All models are wrong – Systems thinking) and at that time did not really know how to fit them into my mental structure for strategic thinking. Only now that I have studied

linearity/ non-linearity, do I have a better understanding of their place and importance in defining non-linear thinking and the implications for strategy.

In this study I wanted to find out what is meant by non-linearity and how you recognise a non-linear effect when you see it. The complexity books that I read, kept referring to power law distributions as the signature feature of non-linear complex adaptive systems. This study is therefore a study of power laws and through studying power laws I want to understand non-linearity.

At first I thought it might be too much of a jump to do a review of power law examples coming from completely different subjects. However, when I was invited by my supervisor Dr Lance Beath to present my research essay (STRA 519) findings to his STRA 502 – Strategic Analysis class in April 2010, I used cross subject examples such as earthquakes, spread of disease and war etc. From the class questions, I soon discovered that strategy/ policy students come from employers such as the Earthquake and War Damage Commission, the Ministry of Health and Defence and what to me may be slightly academic, was to many students part of their day jobs. It is often governmental organisations that have a very high interest and need to understand non-linear systems and where possible, make forecasts and know on what they should be forecasting on.

My inherent focus is on 'strategy', not mathematics/ statistics or modelling

This thesis is inherently about strategy/ policy as it involves forecasting. My interest is not in the statistical or mathematical aspects of power laws. Whilst my thesis is not intended to be a mathematical/ statistics exercise, I was very keen to include the forty or more 'log-log' power law charts into this study. To me these charts are of high importance in showing the actual empirical data that all the subsequent assumptions are based on and also to show how different (as well as the same) each chart is from each other. The reader should note that there are at least two different approaches to drawing these charts (e.g. one is the Zipf's approach of ranking and the second is simply numeric). Seeing the various charts also highlights how the 'real' data is never completely 'perfect', but it is still always amazing to see these straight power law lines considering they are 'emergent' properties.

Whilst I have no inherent interest in producing 'models', I note that 'complexity' and probably every topic covered in my thesis can be and already is being modelled. Power laws have given modellers a wonderful tool to start their work. We need their work, so good luck to them, but that is not me. I have no inherent interest or skill in the mathematical aspects, only in how a policy/ strategy person can use the results for increased forecasting understanding and the understanding that predicting the next event is effectively impossible when dealing with a non-linear effect.

High level starting point

From the conclusions of the research essay (STRA 519) and subsequent readings, I compiled the following linear versus non-linear table which set out my initial subject understanding. In the next three chapters, I will define the various terms I have used below.

Table 1 Comparing non-linear to linear

	Non-linear	**Linear**
Definition	Complex adaptive systems	Simple systems
Change phase	'Punctuated'	'Equilibrium'
Statistical distribution system applying	Power laws	Normal bell curve
Key strategic features	The whole is greater/less than the sum of the parts and the need to leverage	The whole is equal to the sum of the parts

This thesis sets out to explore the validity of this initial categorisation. To do this and to keep tight study focus, I selected as the prime focus for this thesis, the study of **power law effects**. To do this I have to take examples from where-ever I can across subjects. This includes the study of wealth distribution (Pareto's law), the frequency of words and city sizes (Zipf's law), share market fluctuations, company size fluctuations – including growth/ decline rates, company failures, earthquakes (Gutenberg-Richter law), scientific citations (Lotka's law) and scientific progress (Thomas Kuhn), species extinction, war (Richardson's law), river networks, sexual partners, traffic jams and computer networks. Whilst many of these different power effects had their own name (e.g. Pareto's law of wealth distribution), behind each power law is the same defining log-log graph and the same distinctive set of rules.

Of course life is never simple, as my early readings (i.e. Mark Newman) had highlighted that the study of power laws is a relatively new subject and that the defining 'log-log' graphs are not only caused by power law distributions but also log-normal and exponential distribution. I am not sure what the significance is of these other distributions but I have kept my focus on purely power law distributions since the power law distribution system is the basis of complexity and the distinctive scale free predictions. These issues are considered further in Chapter 17.

Thesis outline

Chapters Two to Four define key concepts. Chapter Two starts with understanding non-linearity (complexity); by considering the following distinctions – reductionism versus holism, simple versus complex adaptive systems, linear versus non-linear examples; as well as providing examples of how non-linear systems produce 'emergent' properties or effects. At

the end of this chapter, I make the link between strategic planning/ linearity, and strategic thinking/ non-linearity.

Chapter Three introduces the concept of power laws. This is done by highlighting the difference between the normal distribution and power law distribution systems. This chapter introduces the 'log-log' graph which is the signature of a power law, and describes how power law distributions are 'scale-free'. In this chapter I assert that understanding power laws are fundamental for strategic forecasting.

Chapter Four introduces the concepts of punctuated equilibrium, the 1987 landmark physics 'sandpile' experiment and metaphor and related to this, self organised criticality. The 'sandpile' metaphor is considered in some detail because this is perhaps the key metaphor in all power law examples throughout my thesis. Whilst the term 'punctuated equilibrium' was first used to explain fossils, the term has subsequently been applied to other subjects including strategy (especially by Miller, Friesen, and Mintzberg).

Chapter Five is a review of Pareto's law of wealth distribution. This is the first of the different power law distributions covered. This chapter highlights how classical economics (linear/ equilibrium/ rational) cannot account for the non-linear Pareto's law, and it explores the 80/20 Rule and its conceptual opposite approach of the 'Long Tail'.

Chapters Six and Seven consider different power law applications that are called Zipf's law. These include the frequency of words used, size of cities, and size of companies. All these examples have log-log graphs that provide high empirical evidence of non-linearity and have little traditional linear theory to explain them. The non-linear/ complexity explanations are relatively new and still developing.

Chapters Eight to Eleven consider different economic power law applications. These include the distinction between traditional 'equilibrium' economics and complexity 'non-equilibrium' economics, as well as considering non-linear concepts such as Joseph Schumpeter's 'Gales of creative destruction' and strategy is 'about surviving the Red Queen Race'; as well as considering power law charts for stock price fluctuations, share volumes traded, company growth/ decline rates, and company failure rates.

Chapters Twelve to Sixteen cover a number of different networks, both natural and human that have power law distributions. These include earthquakes, scientific citations, Thomas Kuhn's paradigm shifts, species extinctions, war, movie-actor networks, the Internet, the World Wide Web, traffic jams, river networks, and the number of sexual partners.

Chapters Seventeen to Nineteen consider what this all means. Chapter Seventeen considers the problem, that not all apparent power law distributions are completely power law distributions. Chapter Eighteen considers the strategic implications of non-linearity and Chapter Nineteen offers my conclusions to the question I originally posed. This was, how do you know when a situation is non-linear and what tools and techniques do you use to handle that situation?

Chapter 2 - Understanding non linearity (Complexity)

This chapter considers four important background and interrelated sets of definitions that are important for understanding forecasting when power laws apply. These are:

- Defining reductionist and holistic approaches,
- Defining simple and complex adaptive systems,
- Defining linear and non-linearity, including examples, and
- Defining what is meant by non-linear systems producing 'emergent' properties or patterns.

Defining reductionist and holistic approaches

Historically scientific thinking has always placed a lot of value in 'reductionism', which seeks to break things (e.g. atoms or DNA etc) into smaller and smaller parts. Reductionism asserts that the whole can be decomposed and analyzed in terms of its constituents that are considered as fundamental to the whole.[1] This has generally served science very well and has lead to great advances in most subjects over the last three hundred years.

In recent decades, many scientists (especially those studying complexity) themselves have questioned the value of relying entirely on reductionism and have suggested as a complementary style, the use of 'holistic' thinking. Holism professes that the whole is fundamental and cannot be reduced to its components, as the whole is sometimes greater than the sum of its components.[1] This distinction (reductionism versus holism) has become clearer over time, as we now know that linear systems are more or less equal to the sum of their parts; and nonlinear systems may be either much more, or much less, than the sum of their parts.[2]

Reductionism is great for quantitative aspects of internal details. In contrast, our current understanding of external large-scale effects is mostly descriptive and qualitative, geometric rather than numerical. We can recognise a hurricane from satellite photos, but we can't tell exactly what it's going to do.[3] For many subjects such as in ecology, epidemiology, or economics, we also need to study the whole or at least part of the system. Without the holistic thinking provided by 'complexity' it is difficult to make predictions on the spread of AIDS or to predict stock-market crashes. Cohen goes further and suggests that there are no big areas of reductionist causality in social science or management studies.[4]

---

[1] Thuan, Trinh Xuan (2001). *Chaos and Harmony. Perspectives on Scientific Revolutions of the Twentieth Century.* Oxford University Press. New York. USA. Pp. 340-346.
[2] Gribbin, John (2004). *Deep Simplicity. Chaos, Complexity and the Emergence of Life.* Allen Lane. An imprint of Penguin Books. London, England. Pp. 3-49.
[3] Cohen, Jack and Stewart, Ian (1994). *The Collapse of Chaos. Discovering Simplicity in a Complex World. Viking.* Published by Penguin Group. Great Britain. Pp. 442.
[4] Ibid., Pp. 182.

Defining simple and complex systems

Simple systems generally involve a small number of individual elements with relatively weak interactions between them. Reductionism works well for simple systems.[5] Using this definition, simple systems would include many aspects of classical physics but would be very problematic for much of modern physics and biology.

Complex adaptive systems, on the other hand, involve a medium-sized number of agents (drivers, traders, molecules etc). These agents are generally both intelligent and adaptive, in the sense that they make decisions in accordance with various rules, and are ready to modify their rules of action on the basis of new information that comes their way. There are no dictators or centralised controllers in these systems. There is generally no single driver, trader, or molecule that has access to what other actors in the system are doing, so the agents in a complex adaptive system make their decisions and update their rules for action on the basis of local rather than global, information. For studying complex adaptive systems, new and different approaches to reductionism are needed. Fortunately one new method to study complex systems is by computer simulation and 'what-if' games. No longer do we automatically have to break the system into simpler subsystems.[5]

Defining linear and non-linear

Linear systems are systems in which changes in the initial conditions lead to proportional changes in the final state.[1] Linear systems have been the mainstay of science for more than three hundred years: because one plus one equals two, and we can predict that the total volume of water flowing down a drain is double when a tap drips for twice as long.[6] Linear systems are relatively simple to understand, especially compared to non-linear systems, so it is probably natural that our education system seems to have a bias towards training us to think in a linear fashion.

Nonlinear systems are systems in which changes in initial conditions do not produce proportional changes in the final state.[1] Nonlinear systems do not obey the simple rules. Compare the simple flow of water down a drain (double the time means double the flow at a constant rate) with the complex nonlinear phenomena that regulate the quantity of water in the human body, or the movement of water vapour in the clouds overhead.

---

[5] Casti, John L (1997). *Would-Be Worlds. How Simulation is Changing the frontiers of Science.* John Wiley and Sons, Inc. New York. U.S.A. Pp. Ix-x.
[6] Coveney, Peter and Highfield, Roger (1995). *Frontiers of Complexity. The search for Order in a Chaotic World.* Fawcett Columbine. New York. U.S.A. Pp. 7-10.

Most physical systems are inherently non-linear in nature.[7] Nonlinearity with its irreversibility, characterizes phenomena in every field of science, whether apparently random weather patterns, the outbreak of flu epidemics, or the spread of information and ideas.[6] They show up in phenomena ranging from turbulence over an aircraft wing, to weather, lasers, and the firing of synapses in the brain. Nonlinear systems are so common, and truly linear systems so relatively rare, that the mathematician Ian Stewart thinks it is just as silly for physicists to have a field called 'nonlinear systems', as it would be for zoologists to have a field called 'nonpachydermology' – the study of all nonelephants.[8] These thoughts are effectively rephrasing the earlier comments of the legendary Hungarian-American mathematician John von Neumann who referred to the theory of nonequilibrium systems (effectively the same thing as non-linear systems) as the 'theory of non-elephants', i.e. it covers most things.[9]

Nonlinearity causes small changes on one level of organisation to produce large effects at the same or different level. This is called positive feedback, which turns amplified music into a deafening howl.[6] The phenomenon where a small change in the initial state of a dynamic system can dramatically alter its subsequent evolution is called the **'Butterfly effect'.**[1] Characterised as the so-called butterfly effect: a butterfly flaps its wings over the Amazon rain forest and sets in motion events that lead to a storm over Chicago. The next time the butterfly flaps its wings, however, nothing of meteorological consequence happens.[10] In non-linear systems small inputs can (but not always) lead to dramatically large consequences. This effect has even been applied to our lives, where it has been suggested that when we look back in detail on the major events of our lives, it is not uncommon to be able to identify such seemingly inconsequential random events that led to big changes.[11]

Nonlinear equations are difficult to solve.[7] The study of nonlinear systems languished until the 1960s and 1970s, when a combination of new mathematical tools and computers reopened their study.[8] Only with computer graphics are we really able to explore the complexity that non-linear equations possess, by literally doing a million possible combinations.[6] Cheap computers and plentiful computing power has enabled researchers to explore complexity in new and often unforeseen ways.[12]

Over recent decades, research has increasingly documented that nonlinear behaviour is typical of the real world and this means in a qualitative sense 'getting more than you

---

[7] Wikipedia - *Nonlinear system.* http://en.wikipedia.org/wiki/Nonlinear_system 19/12/2009

[8] Beinhocker, Eric D. (2007). *The Origin of Wealth. Evolution, Complexity, and The Radical Remaking of Economics.* Random House Business Books. London, Great Britain. Pp. 107.

[9] Bak, Per (1997). *How Nature Works – The Science of Self-Organized Criticality.* Oxford University Press. Oxford, U.K. Pp. 26-32.

[10] Lewin, Roger (1992). *Life at the Edge of Chaos.* Pp. 11.

[11] Mlodinow, Leonard. (2009).*The Drunkard's Walk. How Randomness rules our Lives.* Penguin Books. London. UK. Pp. 194.

[12] Beinhocker, Eric D. (2007), op. cit., Pp. Xii.

bargained for' or 'getting far less than you bargained for', unlike linear systems which produce no surprises, i.e. 'you get what you bargained for'.[13]

Nonlinear systems are very sensitive to their starting conditions, so that a tiny difference in the initial 'push' they first receive causes a big difference in where they end up, and there is feedback, so that is how a system effects its own behaviour. Complexity is based on two simple ideas – the sensitivity of a system to its starting conditions and feedback.[2]

In general, nonlinearity produces complex and frequently unexpected results.[6] Most of nature is non-linear and not easily predicted. For example, weather forecasting is the classic example of being unpredictable. Other non-linear systems such as ecosystems, economic entities, developing embryos or the brain, each are examples of complex dynamics that defy mathematical analysis or simulation using linear approaches. A very slight difference in the initial conditions produces very different outcomes, which is the reason for the unpredictability.[10]

Examples explaining different aspects of non-linearity

In his book, the 'Tipping Point', Gladwell describes his childhood struggles to understand things that are not linear and he gives the example of pouring ketchup on his dinner. Like many children encountering this problem for the first time, he assumed that the solution was linear. He thought that with steadily increasing hits on the base of the bottle it would yield steadily increasing amounts of ketchup out the other end. Not so, Gladwell's father said, and he recited a ditty that effectively sums up non-linearity for Gladwell, being: 'tomato ketchup in a bottle – none will come and then the lot'll'.[14]

Linear thinking assumes that large events need large causes, for example, the large mass extinction that killed the dinosaurs was caused by a very large meteorite. However, non-linear thinking points out that large catastrophic events occur as a consequence of the same dynamics that produces small ordinary events. This observation runs counter to the usual way of thinking about large events, in which we look for equally large causes.[9]

Thuan (a physicist) explains what linear behaviour means by using the example of a dry sponge placed under a tap. At first, the weight of the dry sponge increases proportionally to the number of water drops it soaks up. The sponge doubles its weight when it has absorbed twice the number of drops, and triples it with three times as many drops. The relationship between the weight and the number of drops is said to be linear. However, the behaviour of the sponge changes by the time it has soaked up a lot of water. There comes a point when its

---

[13] Coveney, Peter and Highfield, Roger (1995). op. cit., Pp. 430.

[14] Taleb, Nassim Nicholas (2007). *Fooled by Randomness. The Hidden Role of Chance in the Markets.* Penguin Books. London. England. Pp. 282.

ability to absorb extra water diminishes, until it becomes zero. When that happens, the weight of the sponge no longer increases in proportion to the number of drops falling on it. The behaviour switches from linear to non-linear. Finally, the weight of the sponge reaches a steady level, as it cannot absorb any more water. It has reached saturation.[15]

Thuan uses the same saturation metaphor of phase changes when we visit a museum full of works of art. At first, our aesthetical pleasure increases in proportion to the number of paintings we get to admire. We are in the linear phase. After a while, though, the pleasure wears thin; we gradually grow numbed by too many masterpieces, and we enter a non-linear regime. Eventually, we reach a point where we can not stand it any more. We have become saturated with paintings.[15]

Biology and economics have traditionally been thought of as in a balanced equilibrium stable state. In the equilibrium state, small shocks will cause only small disturbances, modifying the equilibrium state only slightly. The system's response is proportional to the size of the impact; equilibrium systems are said to be 'linear'. Small freak events can never have dramatic consequences.[9]

Non-linear systems have 'emergent' properties or patterns

From the 1970s, scientists started to refer to systems that were far from equilibrium, that were dynamic and complex, and that never settled into a state of rest as 'complex systems'. These systems have dynamically interacting parts or particles. In such systems the micro-level interactions of the parts or particles lead to the emergence of coherent macro-level patterns of behaviour.[16] These complex systems are not at rest, as they are continuously evolving.[6]

Emergent properties show that the whole can be more than the sum of the parts[17] and the so-called emergent properties can only be described at higher levels than those of the individual units.[6] The central idea of emergence is the idea that meaningful order can emerge spontaneously, all on its own in complex systems made of numerous interacting parts.[18] An emergent property cannot be deduced or explained in terms of the properties of its constituents. In other words, the whole is greater than the sum of its parts.[1]

Examples of emergent properties include:
- The pattern of the **whirlpool** is the result of the dynamic interactions between the individual water molecules. One cannot have a whirlpool with a single water molecule;

---

[15] Thuan, Trinh Xuan (2001), op. cit., Pp. 308.
[16] Beinhocker, Eric D. (2007), op. cit., Pp.18.
[17] Ball, Philip (2004). *Critical Mass – How one thing leads to another.* Arrow Books. London. United Kingdom. Pp. 155.
[18] Buchanan, Mark (2002). *Nexus – Small Worlds and the Groundbreaking Science of Networks.* W.W. Norton and Company. New York. U.S.A. Pp. 198.

rather, the whirlpool is a collective or '**emergent**' property of the system itself. It is not an expected result, if you just study the properties of a single water molecule sitting in isolation. You need to have a few billion water molecules together, add some energy in the right way, and then you get the complex macro pattern of a whirlpool.[16]

- The interactions of millions of people, making decisions, acting out their cultural norms, cooperating, competing, and going about their daily lives, creates an emergent phenomena that we call '**society**' which has different features than those of the studies of individuals that make it up.[19]

- In the economic world, Adam Smith's invisible hand of 'market forces' and Pareto's law of wealth distribution represent two principle emergent properties.[18]

- Life is an emergent property, as are human beings.[20]

- Zipf's law of cities or words is an emergent property in the sense that it is not an obvious consequence of the underlying dynamical rules.[9]

- Based on these examples, it is probably safe to say every power law effect mentioned in this Thesis is an emergent feature.

Networks are important for complex systems

Networks are an essential ingredient in any complex adaptive system. Without interactions between agents, there can be no complexity. For example, the biological world comprises a massive hierarchy of networks: molecules interact in cells, cells interact in organisms, and organisms interact in ecosystems etc.[21]

The economic world likewise depends on networks. The earth is girdled by many networks, including roads, sewers, water systems, electrical grids, railway tracks, gas lines, radio waves, television signals, and fibre-optic cables. These provide the highways and byways of the matter, energy, and information flowing through the open system of the economy. The economy also contains massively complex virtual networks: people interact in companies, companies interact in markets, and markets interact in the global economy. Just as in biology, the networks of the economic world are arranged in hierarchies of networks within networks.[21]

---

[19] Beinhocker, Eric D. (2007), op. cit., Pp. 450.

[20] Coveney, Peter and Highfield, Roger (1995), op. cit., Pp. 330.
[21] Beinhocker, Eric D. (2007), op. cit., Pp. 141.

Pattern prediction, not precise prediction, for complex systems

The study of complexity has its roots in the classical political economy of the Scottish Enlightenment, and then later developed by the Austrian school of economics, which believed that order in market systems is spontaneous (or emergent) in that it is the result of human action, but not the execution of any human design. Key Austrian economists of this school of thinking were Joseph Schumpeter and Friedrich Hayek.[22]

Hayek dedicated much of his work to the study of complex phenomena. One of Hayek's main contributions to early complexity theory is his distinction between the human capacity to predict the behaviour of simple systems and its capacity to predict the behaviour of complex systems through modelling. He believed that economics and the sciences of complex phenomena in general, which in his view included biology, psychology, and so on, could not be modelled after the sciences that deal with essentially simple phenomena like classical physics. Hayek believed you could only explain complex phenomena through modelling, and this will only give you pattern prediction, compared with the precise predictions that can be made out of simple systems.[22]

Key heroes of non-linearity

It is an interesting observation that most of the people I have named throughout this thesis, such as Schumpeter, Hayek, Mandelbrot, Pareto, Zipf, Taleb and perhaps even Mintzberg were all initially dissenters or heretics to the conventional wisdom of their day. Not only were they dissenters, they were unpopular. Fortunately times change and Schumpeter is now considered equal (or even greater) to John Maynard Keynes, as being one of the greatest economists of the twentieth century. Hayek received a Nobel Prize and most people have a vague understanding of Pareto's 80/20 law.

The following is my summary of key words used to compare linear with non linear effects.

---

[22] Wikipedia – *Complex Systems*.
http://en.wikipedia.org/wiki/Complex_systems#Complexity_and_Chaos_theory 07/05/2010

Table 2

Linear attributes	Non-linear attributes
In equilibrium	Far from equilibrium
Top down	Bottom up
Simple systems	Complex systems
Certainty	Uncertainty
Additive (sum of the parts)	Emergent (whole greater or less than the sum of the parts)
Reductionist	Holistic
Stable	Unstable
Time and nature fixed	One period effects the next
Static	Dynamic
Trajectory	Abrupt
Continuity or evolutionary change	Discontinuous or revolutionary changes
Detail	Big picture
Tidy	Messy
Analysis	Synthetic
Deterministic	Iterative calculations
Solvable - repeatable	Simulation
'Rational'	'Bounded rationality'
Precise predictions	Pattern predictions only
Linear effects	Small causes with tipping points
Basis for Strategic Planning	**Basis for Strategic Thinking**

Summary of key words used to compare linear to non linear effects.

The strategic planning/ strategic thinking distinction could also be categorised as a linear versus non-linear systems debate, or in the wider science/ mathematics debate, strategic thinking fits into non-linear thinking which is called 'complexity'. This is a subject that focuses on those 'abrupt' changes in different physical/ social systems including those in finance and in economics. This means that 'strategy' fits into a bigger intellectual framework that cuts across science, mathematics, economics and finance; since most real world systems are non-linear. Non-linear systems are noted for 'tipping points', power laws and abrupt transformational changes.

Since both linear/ non-linear systems are purely natural, it therefore follows that a strategist needs a good understanding of linearity and non-linearity as well as strategic planning and strategic thinking. I suggest that under a linear system/ phase, the strategic planning approach should be dominant, but supported by strategic thinking (since you never know when events will turn abruptly); whereas in a turbulent non-linear period the strategic thinking approach should be dominant, but supported by strategic planning (since you know that great turbulence will not last). This has a loose parallel in the theory of the left/ right brain split,

where it is not wise to use only one style of thinking, since there are two styles which suit different situations. The key is to pick the right thinking style for the right situation. Just as we have one brain, but two thinking styles, so in the strategy toolbox we also have two valid, useful and complimentary general strategic approaches.

Chapter 3 - Introduction to power laws

The normal versus power law distributions

One type of empirical relation that seems to turn up regularly in complex-systems (non-linear) studies is a power law.[1] Power laws are in my opinion of profound importance in understanding forecasting in non-linear systems. Defining power laws is therefore the subject of this chapter. To understand power laws you first have to understand the 'normal' distribution.

For most of the last two hundred years the 'normal' (also called 'Gaussian' or 'Bell' curve) distribution has been the dominant statistical distribution to work with. Many quantities in nature follow the bell curve. For example, the peaked distribution characterizing random networks, height of adult males, IQ distributions, velocity of molecules in a gas, etc.[2] Normal distributions cluster around a typical value. Examples of this are, the speed of cars on a motorway, the weight of apples in a store, air pressure, sea level, or the temperature in Wellington at noon on Christmas Day. All these things vary, but their distributions strongly cluster around an average point. Few observations are far from the typical value, making the typical value representative of most observations. For instance, it is a useful statement to say that an adult male American is about 180cm tall because no adult male American deviates very far from this size. Even the largest deviations, which are exceptionally rare, are still only about a factor of two from the mean in either direction and hence the distribution can be well-characterized by quoting just its mean and standard deviation.[3]

By contrast, Barabasi believes that power laws are at the heart of some of the most stunning conceptual advances in the second half of the twentieth century, emerging in fields like chaos, fractals, and phase transitions. The fact that the networks behind the Web, Hollywood, scientists, the cell, and many other complex systems all obey a power law allows us to understand Pareto and all other situations when the 80/20 rule applies. This means we can for the first time in history start to understand complex systems.[2]

Power laws are very different from the bell curves describing our height. First, a power law distribution does not have a peak. Rather, a histogram following a power law is a continuously decreasing curve, implying that many small events coexist with a few large events. The distinguishing feature of a power law is not only that there are many small events but that the

---

[1] Casti, John L (1997). *Would-Be Worlds. How Simulation is Changing the frontiers of Science.* John Wiley and Sons, Inc. New York. U.S.A. Pp.126.

[2] Barabasi, Albert-Laszlo (2002). Linked – The New Science of Networks. Perseus Publishing. Cambridge, Massachusetts, USA. Pp. 67-72.

[3] Clauset, Aaron, Shalizi, Cosma Rohilla, and Newman M.E.J. *Power-Law Distributions in Empirical Data.* arXiv:0706.1062v2 (physics.data-an) 2 Feb 2009 www-personal.umich.edu/-mejn/pubs.html (Publications of Mark Newman) University of Michigan. Pp. 1.

numerous tiny events coexist with a few very large ones. These extraordinarily large events are simply outside of the rules of a bell curve.[2]

There is an important qualitative difference between a power law and a bell curve when it comes to the tail of the distribution (Refer Figure 1). Bell curves have an exponentially decaying tail, which is a much faster decrease than that displayed by a power law. Power laws decay far more slowly than the Bell curve, allowing for large 'rare events' that can never happen in a Bell curve. These large 'rare' events are an important characteristic of power laws. Power laws formulate in mathematical terms the notion that a few large events carry most of the action.[2]

Figure 1
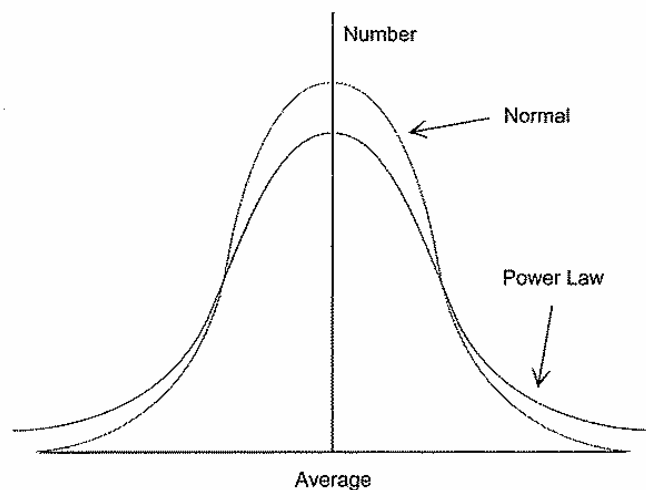
Power law curve compared to a normal distribution curve. (Buchanan – 2007).

Barabasi gives an example of the difference between the normal distribution and a power law distribution. He compares a U.S. roadmap with an airline routing map. On the roadmap cities are the nodes and the highways connecting them the links. This is a fairly uniform grid pattern. Each major city has at least one link to the highway system, and there are no cities served by hundreds of highways. Thus most nodes are fairly similar, with roughly the same number of links. Such uniformity is an inherent property of random networks with a peaked degree (normal) distribution.[2]

The airline routing map differs drastically from the roadmap. The nodes of this network are airports connected by direct flights between them. Inspecting the flight map quickly shows that there are a few important hubs, such as Chicago, Los Angeles, Dallas, Denver, Atlanta, and New York from which flights depart to a large number of other U.S. airports. The vast majority of airports are tiny, appearing as nodes with at most a few links connecting them to one or several hubs. Thus, in contrast to the highway map, where most nodes are equivalent, on an airline map, a few hubs connect hundreds of small airports.[2] Whilst the number of highway

links to each city produces a Bell curve indicating a normal distribution, the number of flight links to each U.S. city produces a power law graph. (Refer Figure 2). This difference is profound and each of the two statistical systems follow different sets of rules. Generally people have good familiarity with the normal distribution but have only superficial knowledge of power laws. For example, most people have heard of the Pareto's 80/20 Rule.

Figure 2
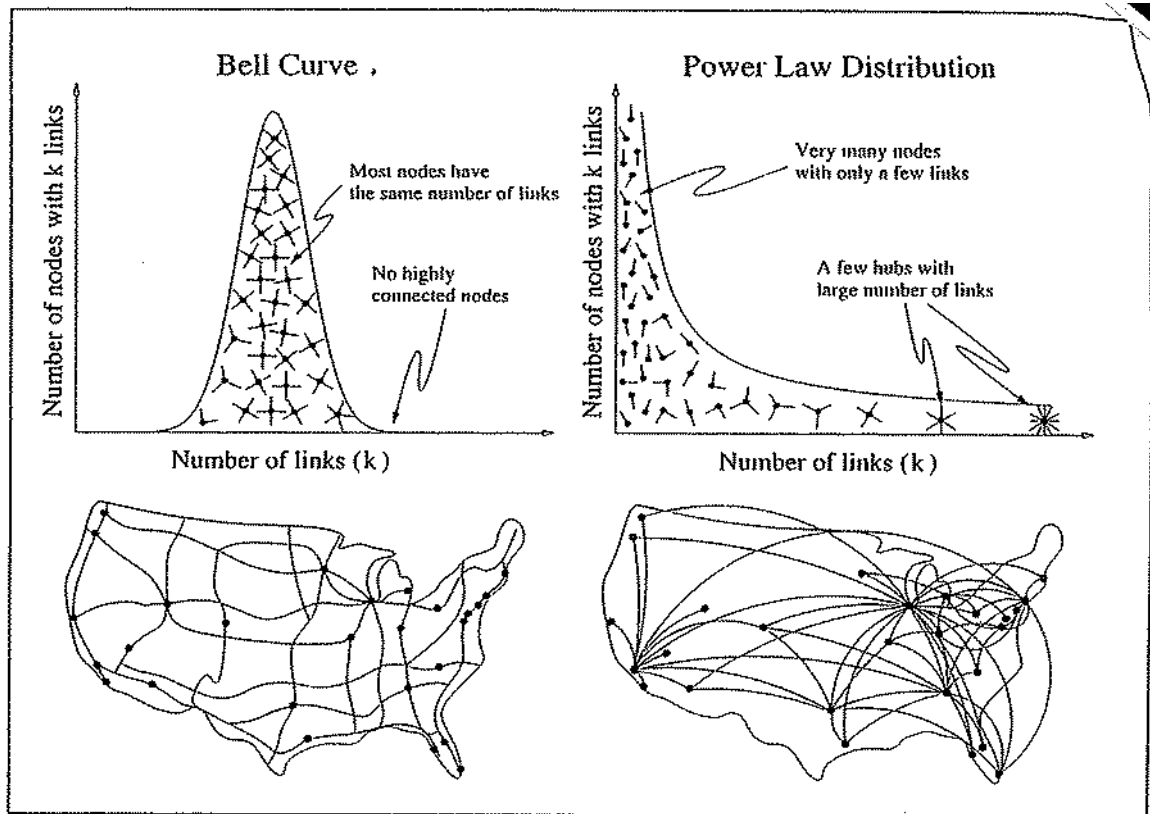
The Random (Normal or Bell Curve) versus the Scale-free Networks (Power law Distribution). (Barabasi - 2002).

The straight line 'log-log' graph is the signature of a power law

All power laws tend to look alike when plotted on a graph with a logarithmic scale. (Where each division is a factor of 10 larger than the one that came before it, for example, 10, 100, 1,000 and so on.) Common examples of logarithmic scales are the Richter scale of earthquakes and the decibel scale of sound volume. When a power law is plotted on a log scale on both axes (called a 'log-log graph'), a straight line sloping down should result. The exact angle of that slope varies, but whether it's the size and frequency of earthquakes or the distribution of publicly listed companies by stock-market capitalization, the slope is always a straight line.[4] (Refer Figure 3).

---

[4] Anderson, Chris (2006). *The Long Tail – Why the Future of Business is Selling Less of More.* Hyperion. New York. U.S.A. Pp. 126-7.

The straight line log-log graph is often called the signature of a power law. However, whilst this is a necessity, there are non-power data sets (such as 'log-normal' or exponential) that mimic this signature behaviour, so care needs to be taken.[5] Throughout my thesis I have therefore included as many of the straight line log-log graphs as I can find to explain the examples I am describing.

Figure 3
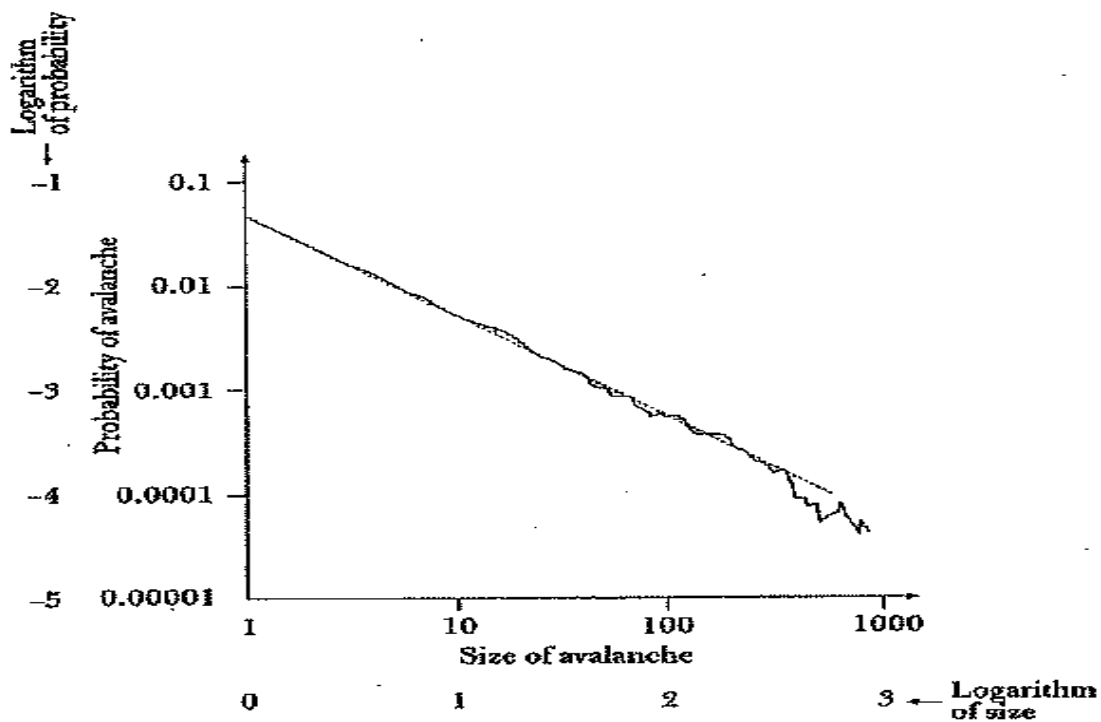
Power law curve for avalanches in a mathematical model of a sand pile. (Ball – 2004).

Only straight line log-log graphs considered in this thesis

For this study I have only used power laws that produce a straight line log-log chart. Therefore Moore's law which show a logarithmic increase in transistor counts over time and Kleiber's law about a power law for metabolic rates compared to animal mass, both produce results that are straight line, but effectively a 45 degree slope from the bottom left hand side and rising on the right hand side. This is a different graph from the graphs in this and other studies of complexity. The slopes in these graphs slope the opposite way – a 45 degree slope from being high on the left hand side and low on the right hand side. Since my desire is to study complexity (non-linear effects), this means that for this study I have looked exclusively at only those activities that produce a straight line log-log chart.

---

[5] Wikipedia – Power Law. http://en.wikipedia.org/wiki/Power_law 07/05/2010

Power law exponents (slope of a straight line log-log graph)

Each power law is characterized by a unique exponent. The power law describes the degree distribution and the exponent is often called the degree exponent. The distribution of incoming links on webpages followed a power law with a degree exponent close to two. The collective action of millions of webpage creators work together in a way that produces this emergent result that is not a 'random' (i.e. not a normal distribution) network.[2]

Taleb gives an example of Pareto's power law of wealth distribution, where if wealth doubles from $1 million to $2 million, the incidence of people with at least that much money is cut in four, which is an exponent of two. If the exponent were one, then the incidence of that wealth or more would be cut in two. The exponent is called the 'power' which is why we use the term power law.[6]

Taleb illustrates how exponents work, when he says that only 96 books in 2004 year sold more than 250,000 copies and that he 'thinks' that the exponent is around 1.5. You can extrapolate to estimate that around 34 books will sell more than 500,000 copies – simply 96 times $(500,000/250,000)^{-1.5}$. Around 12 books should sell more than one million copies, being 96 times $(1,000,000/250,000)^{-1.5}$. The exponent is therefore an important and powerful forecasting tool.[6]

Power law distributions are 'Scale-free'

In a normal distribution, there is an average/ mean and practically all the values are tightly clustered (within three standard deviations) around that figure. In a power law distribution there is no valid average figure that represents the majority of values. The power law average combines the numerous low figures with a few extreme figures and the answer is much greater than the median because of the high outliers. This forces us to abandon the idea of a characteristic node. In a continuous hierarchy there is no single node which we could pick out and claim to be characteristic of all the nodes. There is no intrinsic scale in these networks. This is the reason networks with power-law distributions are called 'scale-free'. Most complex networks in nature have a power-law degree distribution, so are called 'scale-free networks'.[2] This property of 'scale invariance' is one of the most important aspects of power laws and therefore attracts a lot of interest.[5]

Forecasting needs to understand the rare big events or 1/f noise

Power laws are often expressed as the frequency of an event, being equal to 1 divided by some power of its size. Turning this around, the size of an event is proportional to 1 over some power of its frequency. This is called '1 over f noise', and written as 1/f noise. 1/f noise

---

[6] Taleb, Nassim Nicholas (2007). *The Black Swan – The Impact of the Highly Improbable*. Penguin Books. London. England. Pp. 262-5.

turns up in all naturally varying systems and is synonymous to power laws behaviour. For example, music also has 1/f noise, as does the variations in the light from a quasar, as does the weather system, as does earthquakes, which all obey a power law.[7]

1/f noise means that events of any size can happen at any time, but that big events (in this case, big departures from the average) are rare. Traditionally insurance companies and civic planners have worked out the odds against extreme events on a historical basis. They look at how often a particular type and size of event (say, a severe drought) has occurred in the past, and guess that it will happen with the same frequency in the future. They even talk about things like a 'once in a hundred year' drought. But this is only half the story. The 1/f signal tells us that such events actually obey a power law. And that means that if you get a 'once in a hundred year' drought one summer, the odds of getting a second drought of the same severity the following year are exactly the same as if there had not been a drought the previous year. It is not safe to assume that after one 'once in a hundred years' drought it will be a hundred years before the next one happens. That is linear thinking when dealing with a non-linear system.[7]

Equally, though, 1/f noise implies that a record-breaking drought could be followed by a record-breaking freeze. Next year might be the coldest ever recorded, without this in any way invalidating the greenhouse effect hypothesis, thanks to a random fluctuation of the natural weather system. That is why it is important to look at long-term trends, not just one-off events.[7]

Beinhocker believes that the combination of punctuated equilibrium and power laws in complex adaptive systems is almost fiendishly designed to lull people into a sense of self-confidence, and then pull a nasty surprise. Humans are superb pattern-recognizers. We tend to make predictions by looking for patterns in the past and then extrapolating those patterns out into the future. More often than not, this works quite well. Most of the time, the world is fairly stable. Quite often, conventional strategy analysis can tell us very useful things, and we are able to make accurate-enough short term predictions. But the frequent occurrence of 'stable' periods between punctuations points only gives us a false sense of predictability – there is nothing more dangerous than a 'stable' industry. Likewise, the existence of power laws means that big changes are far more likely to occur than the expectations our mental models give us on the basis of past experience.[8]

As Anderson points out, power-law distributions in social phenomena destroy the idea that what matters is the 'average' behaviour, or the 'average' man/ person. In effect, a power law

[7] Gribbin, John (2004). *Deep Simplicity. Chaos, Complexity and the Emergence of Life.* Allen Lane. An imprint of Penguin Books. London, England. Pp.144-7.

[8] Beinhocker, Eric D. (2007). *The Origin of Wealth. Evolution, Complexity, and The Radical Remaking of Economics.* Random House Business Books. London, Great Britain. Pp. 328.

raises the odds of extreme events, which a Gaussian probability distribution relegates to the status of negligible aberrations.[9]

Understanding power laws are fundamental for strategic forecasting

In my view, understanding 'power laws' is fundamental for strategic forecasting. First it is important to understand that there are two different and completely valid statistical distribution systems operating in the world. The first is the Gaussian or normal distribution system and the second is the more recently defined power laws with all the different names and effects as described later in my Thesis. As Mandelbrot points out, power laws may in the future come to be seen as the 'norm' for most forecasting applications, but the name 'normal' distribution was given to the Gaussian distribution since that was developed first and that distribution type was considered at that time to be the 'norm'.

Not only does the person doing strategic forecasting need to understand at a high level the difference between the two distribution systems, but you also need to be crystal clear as to which distribution system is operating in the environment that you are working in. If risks and likely outcomes of the strategic forecasting are affected by power law distributions then this has profound implications, as all potential events are not equal when power laws are involved. A few events have a massive influence, whilst most will be minor. Most of the standard strategic tools I have used such as SWOT, TOWS or brainstorming assume that all strategies are equal. Under a power laws environment, the risks/ rewards distribution is extremely unfair. The results will be logarithmic, i.e. 1, 10, 100, with most being small, a few being large and fewer still being very large. It is the few very large events that are important in power law distributions and they are the defining feature that is different to the normal distribution.

---

[9] Ball, Philip (2004). *Critical Mass – How one thing leads to another*. Arrow Books. London. United Kingdom. Pp. 307-8.

Chapter 4 - Punctuated equilibrium, 'sandpiles' and self organised criticality

Punctuated Equilibrium – evolution and fossils

Since Darwin, most biologists assumed that evolution proceeded in a stately and relatively linear fashion, leading to a smooth pattern of speciation and extinction. Then in 1972, the palaeontologists Niles Eldredge and Stephen Jay Gould overturned this conventional wisdom and argued that the fossil record shows that biological evolution has not followed a smooth path. Rather, they argued that it has gone through long periods of relatively little change, which they called 'stasis', interspersed with periods of explosive innovation and periods of massive extinction.[1] Gould and Eldredge pointed out that the apparent equilibrium is a period of tranquillity, or stasis, between intermittent bursts of activity and volatility in which many species become extinct and new ones emerge.[2] In summary, the great majority of species originate and become extinct in geological moments (punctuations) and persist in stasis.[3]

For example, during the Cambrian period about 550 million years ago, a burst of evolutionary innovation saw the takeover of the earth by multicellular life and the creation of the forerunners of the species on earth today. Then, about 245 million years ago, during the late Permian period, there was what Gould called 'the grand-daddy of all extinctions', when 96 percent of all marine species on earth disappeared.[1] Gould found that many ancient species and ecosystems often remain stable in the fossil record for millions of years and then either die out or transform themselves into something new in a geological instant.[4]

Gould coined the term 'punctuated equilibrium' to describe this pattern of alternating calm and storm, or periods of quiescence or stasis interspersed with periods of change.[1] He noted that one species does not gradually transform into another, rather it happens fast. Punctuated equilibrium is contrasted with gradualism, which states that evolution generally occurs by smooth steady and continuous progress.[5]

Punctuated equilibrium – a 'signature' of complex systems

Patterns of punctuated equilibrium show up not just in biological evolution, but in other complex systems ranging from physical avalanches, the crashes of stock markets[1], to

---

[1] Beinhocker, Eric D. (2007). *The Origin of Wealth. Evolution, Complexity, and The Radical Remaking of Economics.* Random House Business Books. London, Great Britain. Pp. 173.

[2] Bak, Per (1997). *How Nature Works – The Science of Self-Organized Criticality*. Oxford University Press. Oxford, U.K. Pp. 28-29.

[3] Gould, Stephen Jay (2007). *Punctuated Equilibrium.* The Belknap Press of Harvard University Press. Cambridge, Massachusetts. U.S.A. Back cover.

[4] Waldrop, M. Mitchell (1992). *Complexity – The Emerging Science of Order and Chaos*. Simon and Schuster. New York. U.S.A. Pp. 9.

[5] Wikipedia – *Punctuated Equilibrium.* http://en.wikipedia.org/wiki/Punctuated_equilibrium Retrieved 07/08/2010.

changes of political regimes, such as in Egypt with the abrupt and unexpected fall from power of President Hosni Mubarak after 30 years of rule. The current wave of revolution and unrest sweeping the Middle East shows clear signs of 'punctuated equilibrium'. Per Bak (a Danish Physicist) wrote that, 'Punctuated equilibrium takes place and is at the heart of the dynamics of complex systems. Large intermittent bursts have no place in equilibrium systems, but are ubiquitous in history, biology and economics'.[2]

Gould believed that the two most overarching theories of punctuated change are Kuhn's for human thought and Prigogine's for the natural world. During equilibrium periods, systems are stable, make adjustments that preserve the deep structure against internal and external pressures, and move incrementally. Revolutions are relatively brief periods and vary in magnitude.[6]

Kuhn's notion of the history of change in scientific concept advanced a punctuated theory for the history of ideas – going from stable 'paradigms' of 'normal science' in the 'puzzle solving' mode, through accumulating anomalies that build anxiety but do not yet force the basic structure to change, through rapid transitions to new paradigms so different from the old that even 'conserved' technical terms change their meaning to a sufficient extent that the two successive theories become 'incommensurable'. Kuhn's work on punctuated thinking had a big effect on other disciplines, not least because Kuhn himself said that historians of literature, of music, of the arts, of political development, and many other human activities have long described their subjects in the same way.[6]

Gould pointed out that punctuated equilibrium has been applied to personal growth and that many people go through the greatest changes in their lives in the shortest chunks of time. For example, someone who is described as having living through years of stagnation, raced through a decade of personal growth in the first year of a new career.[6]

In economics, punctuated equilibrium describes the episodic change and long plateaus in the history of markets and prices.[6]

The 'punctuated equilibrium' theory of strategy

Miller and Friesen proposed the '**punctuated equilibrium**' theory of strategy, where there are long periods of evolution and incremental adaptation in traditional directions (more of the same); that are interrupted or punctuated occasionally by short bursts of 'quantum' revolutionary realignment, that are sudden, dramatic, and pervasive shifts in direction.[7]

---

[6] Gould, Stephen Jay (2007), op. cit., Pp. 274-299.

[7] Mintzberg, Henry (2007). *Tracking Strategies – Toward a General Theory*. Oxford University Press. Great Britain. Pp. 202 and 303.

Mintzberg notes that once this sudden and turbulent environment happens, even the very best planning techniques are of no use because of the impossibility of predicting the kind of stability that will eventually emerge. In response to this kind of inconsistency in the environment, patterns of strategic change are never steady, but rather irregular and ad hoc, with a complex intermingling of periods of continuity, change, flux, limbo, and so on. Strategies do not commonly change in continuous incremental fashion; rather, change, even incremental change – takes place in spurts, each followed by a period of continuity. This hints that strategies come in all sizes, as does change.[8]

Strategies that may have been bold and innovative at the outset (as a means to find and develop a niche) tend to become incremental and conservative in the mature years (to protect the niche).[9] Strategic reorientations really are cultural revolutions.[10]

'Sandpiles' and self organised criticality

Self-organised criticality (SOC) has been described as one of a number of important discoveries made in statistical physics and related fields over the latter half of the 20th century, discoveries which relate particularly to the study of complexity in nature.[11]

The 1987 landmark 'sand pile' experiment and metaphor

Physics has long had different metaphors to explain reality. During most of the twentieth century the image of 'billiard balls' was often used to describe atoms etc. This changed in 1987 when three theoretical physicists Per Bak, Chao Tang and Kurt Weisenfeld carried out what today is a landmark experiment. They carried out a 'sandpile' experiment, where they were trying to imagine what would happen if someone were to sprinkle grains of sand one at a time onto a table top. As grains pile up, it seems clear that a broad mountain of sand should edge slowly skywards. As the pile grows, its sides become steeper, and it becomes more likely that the next falling grain will trigger an avalanche. Sand would then slide downhill to some flatter region below, making the mountain smaller, not bigger. As a result, the mountain should alternately grow and shrink, its jagged silhouette forever fluctuating.[12]

Bak and his colleagues used a computer to drop these imaginary 'grains' onto an imaginary 'table', with a few simple rules dictating how grains would topple downhill as the pile grew steeper. It was not quite the same as a real sand pile, and yet the computer had one spectacular advantage – a pile would grow in seconds rather than in days. They were able to plot the results after they ran a huge number of tests, counting the grains in millions of

---

[8] Ibid., Pp. 27-28.

[9] Ibid., Pp. 174.

[10] Henry Mintzberg (1996) *Crafting Strategy*. Reading 5.2 in the *Strategy Process: Concepts, Contexts, Cases.* Prentice Hall, New Jersey. Pp 106-9.

[11] Wikipedia – Self Organised criticality. http://en.wikipedia.org/wiki/Self-organised_criticality 07/05/2010

[12] Buchanan, Mark (2000). *Ubiquity – The Science of history…Or Why the World is Simpler Than We Think.* Weidenfeld and Nicholson. London. England. Pp. 11-16.

avalanches in thousands of sand piles. They found a surprising result; that there was no typical size of an avalanche. Some involved a single grain; others ten, a hundred, or a thousand. Still others were pile-wide cataclysms involving millions that brought nearly the whole mountain tumbling down. At any time, literally anything, it seemed, might be just about to happen.[12]

The next experiment Bak and his colleagues did was to change the colour of the grains from green to red once the sand pile had started to get steep. With more grains, the scattering of red danger spots grew until a dense skeleton of instability ran through the pile. When the red spots come to blanket the pile, the consequences of the next grain became fiendishly unpredictable. It might trigger only a few tumblings, or it might instead set off a cataclysmic chain reaction involving millions. This hypersensitive state to which the computer sand pile organises itself is known as the ***critical state***.[12]

Each avalanche releases 'tension' in the pile, lowering the angle of the slope and restoring stability. But only just. The avalanche merely returns the pile to the brink of sliding, so that the very next grain that it adds could trigger another landslide. So the sand pile hovers around this state of precarious stability, never deviating far from it. Rather than the pile being liable to collapse irretrievably, the critical state is constantly resurrected after each avalanche. This is why the **critical state** is said to be **self-organised**. The sand-pile model describes a non-equilibrium situation. It is not unchanging, new grains are constantly being added. This is the driving force that prevents equilibrium from being attained. Self-organized criticality is a property of most non-equilibrium systems, especially where power laws apply.[13]

Bak and his colleagues measured the size and frequency of avalanches and found that they followed a power law probability resulting in the slope of the log v log graph being a straight line. The slope is the exponent of the power law and was found to be close to -1, which is characteristic of self-organising criticality.[13] (Refer Figure 3). There was a power-law distribution of event sizes where there is no characteristic size for events, since the number of events of a certain size is inversely proportional to some power of that size; and 1/f noise applies, where events are distributed over all time scales, but the power or size of events is inversely proportional to some power of their frequency.[14]

The basic notion of criticality has been familiar to physicists for more than a century, yet has always been seen as a kind of theoretical freak, an extremely unstable and unusual condition that arises only under the most exceptional circumstances. In the sand pile, however, it

[13] Ball, Philip (2004). *Critical Mass – How one thing leads to another*. Arrow Books. London. United Kingdom. Pp. 283-299.

[14] Gould, Stephen Jay (2007). *Punctuated Equilibrium.* The Belknap Press of Harvard University Press. Cambridge, Massachusetts. U.S.A. Pp. 234.

seemed to arise naturally and inevitably through the sprinkling of grains. This led Bak and his colleagues to ponder the provocative possibility: if the critical state arises so easily and inevitably in a sand pile, might something like it also arise elsewhere? Could lines of instability of a logically equivalent sort, run through the earth's crust, for example, through forests and ecosystems, and even economies? Subsequent research has shown that the peculiar and exceptionally unstable organisation of the critical state does indeed seem to be ubiquitous in our world.[12]

Key implications of the 'sandpile' experiment include:

- The simple Bak model provides a very good description of the real world and is a very valuable model and metaphor.[15]
- The model provides a mechanism by which the emergence of complexity arises spontaneously from simple local interactions.[11]
- Interesting things happen at the edge of chaos (**critical state**) and feedback is an essential ingredient in what makes them interesting.[15]
- The model shows that the complexity observed **emerged** in a robust manner that did not depend on fine-tuned details of the system.[11]
- The model correctly predicted that the number of grains in each avalanche follows a **power law** and therefore at self-organised critical state, avalanches have a **scale-free or scale invariant** behaviour (any size, any time).[15]
- The model show that when a system is near its critical state, it becomes extremely responsive to disturbances (**Butterfly effect**).[13]
- When the sandpile is in a self-organized critical state, it experiences sudden landslides, interspersed by longer intervals of relative calm, while the sand grains are building up (**Punctuated equilibrium effect**).[15]
- Self organised-criticality has been suggested as the dynamic mechanism behind the theory of 'punctuated equilibrium'.[11]

There are critical points for non-linear systems everywhere. For example they have been found for earthquakes, forest fires, volcanic activity, solar flares, the pattern of species extinctions in the fossil records, fluctuations in economic systems such as financial markets, landscape formation, landslides, epidemics, even in hospital waiting lists and world wars. The characteristics of non-equilibrium self-organized critical-point phenomena, such as extremely sensitivity to fluctuations, 'scale-free' events and, in particular, power law probability distributions are found widely both in nature and in the world of human affairs. Power-laws are the fingerprint of self-organized criticality.[13]

---

[15] Gribbin, John (2004). *Deep Simplicity. Chaos, Complexity and the Emergence of Life.* Allen Lane. An imprint of Penguin Books. London, England. Pp. 172-196.

The critical thing is that there's no telling, when we add a new grain to a pile poised on the brink of an avalanche, whether it will set just a few other grains in motion, or trigger the collapse of the whole slope, or something in between. Big avalanches are less frequent than little ones, but avalanches of all sizes are possible. The fluctuations of the pile, in other words, are scale free.[13]

Self-organizing criticality only tells you about the overall statistics of avalanches; it tells you nothing about any particular avalanche. This is another case where understanding is not the same thing as prediction.[16] For example, knowing the history of events in non-linear systems often provides little or no prediction value as to the next event, despite the fact that each step follows logically from the previous step (historical explanations looking backwards).[17]

Bak believes that self-organized criticality is a new way of viewing nature (non-linear aspects). The basic picture is one where nature is perpetually out of balance, but organized in a poised state - the critical state – where anything can happen within well-defined statistical laws.[18] Self-organised criticality is the way nature is able to make enormous transformations over short time scales.[17]

The principle of Universality

Power laws are not sensitive to particular details. This is the principle of universality, which Per Bak regards as the theorist's dream come true.[19] Universality becomes the guiding principle for understanding many disparate phenomena. It tells us that the laws governing complex systems are simple, reproducible, and ubiquitous.[20] There are many situations around critical points that are independent of the details of the underlying dynamical system. The exponent at the critical point may be the same for many systems in the same group, even though many aspects of the system are different. This makes universality very interesting to physicists.[21]

The Complexity physicists invade other subjects

Buchanan believes that the Per Bak 'critical state' sandpile model has provided some of the most provocative and powerful ideas in all of modern physics. Whilst little known outside physics, these ideas have enabled physicists to establish permanent encampments within

---

[16] Waldrop, M. Mitchell (1992). *Complexity – The Emerging Science of Order and Chaos*. Simon and Schuster. New York. U.S.A. Pp. 306.

[17] Bak, Per (1997). *How Nature Works – The Science of Self-Organized Criticality*. Oxford University Press. Oxford, U.K. Pp. 60-61.

[18] Ibid., Pp. xi.

[19] Ibid., Pp. 42.

[20] Barabasi, Albert-Laszlo (2002). Linked – The New Science of Networks. Perseus Publishing. Cambridge, Massachusetts, USA. Pp. 243.

[21] Taleb, Nassim Nicholas (2007). *The Black Swan – The Impact of the Highly Improbable*. Penguin Books. London. England. Pp. 267.

other subjects such as earthquake science, human physiology, evolutionary biology and economics.[22]

In my thesis, many of the books and articles I have used as my source material come from physicists such as Per Bak, Murray Gell-Mann, Eugene (Gene) Stanley and Mark Newman. Once they started to understand the implications of complexity and power law distributions, they set about with great industriousness looking at other subjects that showed power law effects, for example earthquakes, fossils, share markets, sexual partners, war etc. Initially it seems that subject experts in those other areas were not that excited with the invasion of physicists into their specialist areas. The fact that the 'complexity' thinkers said that details don't count for the big picture, probably made it even worse for the subject specialists who were used to getting donkey deep into the details of their respective subjects. Complexity was really a dream come true for the cross subject theorist of 'universality'.

Stephen J. Gould (Palaeontologist) wrote in one of his books about his reaction to being told by the Danish Physicist Per Bak that he (Gould) was wrong thinking that the extinction shocks had to be external. Per Bak himself writes about how 'flawed' economists are since the traditional economic model believes in 'rationality'. Bak pointed out that since he doesn't claim to be completely 'rational' as defined by the economists' assumptions, how can any 'rational' economic model account for him (or anyone else). Of course once the physicists realised that economics downplays empirical evidence and prides mathematical argumentation and logic, this only increased the cross subject discussion. Bak and others struggled with the idea that economics wasn't as empirically based as other subjects like physics. The physicists seriously questioned the lack of empirical evidence to the assumptions being made in economics. Fortunately some of the great economists like Hayek and Schumpeter had already thought about these issues fifty years ago.

The following chapters highlight these continuing differences in thinking. It is perhaps ironic that the various power law examples I use, all have solid empirical support and little developed theory, whereas classic economic theory is the reverse, elaborate theory that can not explain the various power law effects.

---

[22] Buchanan, Mark (2000), op. cit., Pp. 70.

Chapter 5 - Pareto's law

Pareto's law of wealth distribution

Vilfredo Pareto was an avid gardener, where he noticed that 80% of his peas were produced by only 20 percent of the peapods. He also observed the economic inequality of where 80% of Italy's land and wealth was owned by only 20% of the population.[1] Pareto in effect found power laws (named later by others) when he studied wealth distributions. He observed that this income distribution held true for many countries, regardless of their political system or taxation regime.[2] Pareto found in each country he tested that a small fraction of the population always possesses a lion's share of that country's riches.[3]

Whilst income distributions follow a power law distribution, each set of data will vary around the 80:20 ratio (so the 80:20 is more a rule of thumb) and the log-log graphs may not be a completely straight line and the slope will vary. There seems good reason to believe that a power-law distribution does hold at least for the highest-earning bracket of any society as shown by the straight line in Figure 4. The steepness of such a power-law slope reflects the degree of disparity between incomes: the steeper the slope, the more unequal the distribution of wealth and the more poverty there is, i.e. Haiti or Zaire.[2]

Figure 4
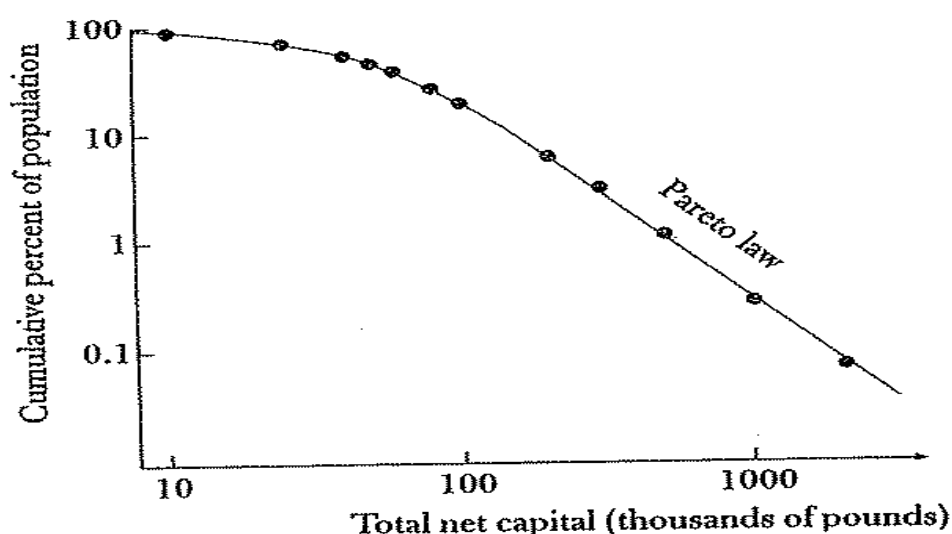
Power law curve for economic wealth (Pareto's law). – The wealth distribution for the population of the U.K. in 1996, according to figures collected by the Inland Revenue. This is a cumulative distribution; each data point shows the percentage of the population with net capital wealth greater than the corresponding amount on the horizontal axis. (Ball – 2004).

[1] Barabasi, Albert-Laszlo (2002). *Linked – The New Science of Networks*. Perseus Publishing. Cambridge, Massachusetts, USA. Pp. 66.

[2] Ball, Philip (2004). *Critical Mass – How one thing leads to another*. Arrow Books. London. United Kingdom. Pp. 307-8.

[3] Buchanan, Mark (2007). *The Social Atom – Why the Rich get Richer, Cheaters Get Caught, and Your Neighbour Usually Looks Like You*. Bloomsbury. New York. U.S.A. Pp. 168-9.

There is nothing special about the numbers 80/ 20, since the Pareto distribution applies to itself. For example; just as 80 percent of effects come from 20 percent of causes, it is also true that 80 percent of that top 80 percent of effects come from 20 percent of that top 20 percent of causes, and so on (i.e. 80% of 80% is 64%; 20% of 20% is 4%, so this implies a 64-4 law is just as valid).[4]

Examples of this uneven wealth distribution (Pareto's law)

- The world's richest 2 percent own more than 50 percent of the world's wealth, whilst the poorest 50 percent of people own 1 percent.[2]

- The income of the 225 richest people in the world is equal to that of the poorest 2.7 billion, 40 percent of the world.[5]

- In New Zealand the top 10% of the population holds 50% of the total wealth, and the bottom half owns around 5%. The top 1% owns about 16% of New Zealand's wealth. This compares with other countries like the US, UK, Australia and Canada where the bottom 30% of the population have less than 1% of the wealth.[6]

- In 1998, according to the Federal Reserve Board Survey of Consumer Finances, the richest 1 percent of U.S. households owned 38 percent of all wealth, and the top 5 percent own more than half of all wealth. The imbalance is greater than in any other advanced industrial country.[7]

- In the United States only 20 percent of the population of the country own 85 per cent of the wealth, and the numbers are similar in Chile, Bolivia, Japan, South Africa, or the nations of Western Europe. It might be 10 percent owning 90 per cent, 5 per cent owning 85 per cent, or 3 per cent owning 96 per cent, but in all cases, wealth seems to migrate into the hands of the few.[3]

- Numerous studies show that wealth inequality all over the world has grown significantly since the early 1980s, in developing countries such as Argentina, China, Pakistan, and South Africa as well as in industrialised ones such as Australia, Finland, Britain, and the United States.[7]

---

[4] Wikipedia – *Pareto principle*. http://en.wikipedia.org/wiki/80/20_rule 04/05/2009

[5] Glenn, Jerome C. and Gordon, Theodore J. and Florescu, Elizabeth. *2007 State of the Future*. Millennium-project.org Accessed 1 September 2009. Pp. 1-5.

[6] Bennett, Neville (2009). *State of the Nation's Wealth.* National Business Review. Pp. 10-12 Wealth Guide.

[7] Buchanan, Mark (2007), op. cit., Pp. Pp. 230.

Economists cannot explain Pareto's law

Pareto's law of wealth distribution stirs strong emotions amongst economists. Of the central issues in economics, as John Kenneth Galbraith wrote, the first is 'how equitable or inequitable is the income distribution. The explanation and rationalization of the resulting inequality has commanded some of the greatest, or in any case some of the most ingenious, talent in the economic profession'. From a mathematical point of view, however, Pareto's law has stubbornly defied explanation. Explanations such as a few people are exceptionally more creative and talented than most, first require an explanation of the inherent spread in human talents and abilities – and that is a tall order. Details such as inheritance, education, desire, luck etc may be part of the factors, but they are the unnecessary detail.[8] What is agreed, is that there is no generally accepted economic explanation for Pareto's law.[3]

The complexity explanation for Pareto's law

The skewed distribution is an **emergent property** of the system. It is a macro behaviour that emerges out of the collective micro behaviour of the population of agents. The combination of the genetic endowments of the agents, where they were born, the rules they follow, the dynamics of their interactions with each other and with their environment, and above all, luck together conspire to give the emergent result of a skewed wealth distribution.[9]

Pareto's law is not about individuals. Rather, it captures a pattern that emerges at the level of many individuals and leaves individual histories aside. Pareto's law reflects little about human cultural, behavioural, or intellectual features and arises instead as a consequence of some deeper principle of organisation.[8]

The number of people having wealth, W, is inversely proportional to W^{α}, where α is a number around two and one-half. In other words, the number of people falls off as the wealth rises to a power between two and three. Each time wealth goes up by a factor of ten, the number of people having that much falls by a factor of about six (about six times as many people have $1 million, for example, as have $10 million). The numbers dwindle in a natural and regular way.[3]

Because of the scale-invariant nature of the power law relationship, the relationship applies also to subsets of the income range. Even if you take the ten wealthiest individuals in the world, we can see the top three (Warren Buffett, Carlos Slim Helu, and Bill Gates) own as much as the next seven put together.[4]

---

[8] Buchanan, Mark (2002). *Nexus – Small Worlds and the Groundbreaking Science of Networks.* W.W. Norton and Company. New York. U.S.A. P189 -190

[9] Beinhocker, Eric D. (2007). *The Origin of Wealth. Evolution, Complexity, and The Radical Remaking of Economics.* Random House Business Books. London, Great Britain. Pp. 86.

Pareto's law of incomes holds for French towns in the nineteenth century, U.S. incomes in 1916, and a number of other populations. The α value does vary between countries. Using current U.S. figures of minimum and maximum incomes – say \$5,000 and \$10^9 Anderson found that half of the wealth is with only $1/8^{th}$ the people. The rich have all the money. The policy implication is that it pays to tax the rich and not the poor. In a failed state such as Haiti or Zaire, α = 1 (for α to be less than 1 is impossible as everyone starves).[10]

Power laws means the extreme is the norm, i.e. 'the winner takes all'

The Pareto Principle is an illustration of a 'power law' relationship. Because it is self-similar over a range of magnitudes, it produces outcomes completely different from Gaussian distribution phenomena.[11] Under a normal or Gaussian distribution, most observations hover around the mean, while the odds of a deviation decline faster and faster (exponentially) as you move away from the mean. 68.2 percent of the observations fall between minus one and plus one standard deviations away from the average. Outliers are increasingly unlikely and you can safely ignore them.[12]

Since much of the real world is controlled as much by the 'tails' of distribution as by 'means' or 'averages'; by the exceptional, not the mean; by the catastrophe, not the steady drip; by the very rich, not the 'middle class', we need to free ourselves from 'average' thinking. Power laws by definition do not follow a normal or Gaussian distribution.[10]

Taleb illustrates the difference between the Gaussian and power law distributions with an example of book sales. He says that if he told you that two authors sold a total of a million copies of their books, the most likely combination is 993,000 copies sold for one and 7,000 copies for the other. This is far more likely than the books each sold 500,000 copies. For any large total, the breakdown will be more and more asymmetric.[12]

The Pareto Principle has been used to attribute the widening economic inequality in the U.S.A to 'skill-biased technical change' of those taking advantage of new technology and globalisation. However, Paul Krugman in the New York Times (February 27, 2006) dismissed this '80-20 fallacy' as being cited 'not because it's true, but because it's comforting'. He asserts that the benefits of economic growth over the last 30 years have largely been

---

10 Anderson, Philip W. (1997) *Some Thoughts about Distribution in Economics* – Part of *The Economy as an Evolving Complex System II.* (Editors Arthur, W. Brian; Durlauf, Steven N.; Lane, David A) - Proceedings Volume XXVII. Santa Fe Institute. Studies in the Sciences of Complexity. Addison-Wesley. The Advanced Book Program. Reading Massachusetts. Pp. 565-6

11 Wikipedia, the free encyclopedia. *Pareto principle.* http://en.wikipedia.org/wiki/80/20_rule Extracted 4 May 2009.

12 Taleb, Nassim Nicholas (2007). *The Black Swan. The Impact of the Highly Improbable.* Random House. New York. Pp. 234-250.

concentrated in the top 1%, rather than the top 20%.[4] This is a clear example of what is called the Matthew effect, where the 'rich get richer'.

Taleb like Krugman, also does not like the phrase '80/20 rule' developed by Pareto, and he also believes that reality is more extreme than that. For example, he believes that the 80/20 rule applies to itself so gets to 64/4 (as mentioned above) and after another iteration will yield approximately 50/1. Taleb says that instead of saying that 80% of the work comes from 20% of the workers, under the 50/1 rule, it is, 50% of the work comes from 1% of the work force. Whilst it may seem unfair and extreme, this seems closer to reality for many things.[12]

Taleb uses the U.S. book business to show how unfair life actually is. The 80/20 rule is only metaphorical; it is not a rule, even less a rigid law. The proportions are more like 97/20 (i.e. 97 percent of book sales are made by 20% of the authors) and it is worse for non fiction when 20 books from approximately eight thousand represent half the sales.[12] Of course for book sales no one knows before hand which books will be in that elite group of best sellers. For simple systems it is sometimes possible to determine before hand the top 20 per cent, but when power laws apply (Complex systems), the reality is that determining the top 20 per cent is only possible after the event.

Hamel gives another example of this 'winner takes all' effect. At the end of 1999, just 32 companies accounted for half the market capitalisation of the S&P 500. This is less than one percent of the companies having 50 percent of market capitalisation.[13]

Pareto distribution – the 80/20 Rule

Pareto's Law or Principle, known also as the 80/20 rule, has been turned into the Murphy's Law of management: 80% of profits are produced by only 20% of the employees, 80% of customer service problems are created by only 20% of consumers, 80% of decisions are made during 20% of meeting time, and so on. It has morphed into a wide range of other truisms as well. For example, 80% of crime is committed by 20% of criminals, 80% of links on the Web point to only 15% of Webpages, 80% of citations go to only 38% of scientists, 80% of links in Hollywood are connected to 30% of actors.[1]

Though it is tempting to infer that the 80/20 rule applies to just about anything, that would be a gross overstatement. In reality all systems following Pareto's Law are a bit special. What sets them apart is a property (power laws) that plays a key role in understanding complex networks as well.[1]

---

[13]   Hamel, Gary (2000). *Leading the Revolution*. Ch.9: The New Innovation Solution. Harvard Business School Press, Boston. Pp. 307.

The Pareto distribution is a power law probability distribution that coincides with many social, scientific, geophysical, actuarial, and other types of observable phenomena.[14] The following examples approximate to Pareto-distributed:

- The sizes of human settlements (few cities, many hamlets/ villages)
- File size distribution of internet traffic (many smaller files, few larger ones)
- The values of oil reserves in oil fields (a few large fields, many small fields)
- The length distribution in jobs assigned supercomputers (a few large ones, many small ones)
- The standardized price returns on individual stocks
- Sizes of sand particles
- Sizes of meteorites
- Number of species per genus
- Areas burnt in forest fires
- Severity of large casualty losses for certain lines of business.[14]

Simplification of thinking by the incorrect use of the 80/20 Rule

Martin is clear about using the 80/20 Rule incorrectly. He believes that conventional or linear thinking tends to take a narrow and simplistic view of causality. We rationalize this simplification by referring to the 80-20 rule. The 80-20 rule implicitly acknowledges that simplification is not the perfect solution to the problems of ambiguity and casual inconsistency, but rather a coping mechanism. We settle for 80 percent to avoid being overwhelmed by complexity. As comforting as simplification can be, however, it impairs every step of the integrative thinking process. It encourages us to edit out salient features rather than consider the question of salience broadly. Simplification makes us favour linear, unidirectional casual relationships, even if reality is more complex and multidirectional. Simplification also encourages us to construct a limited model of the problem before us and to reduce alternatives, giving us the best bad choice available.[15]

Criticisms of the 80/20 Rule

Chris Anderson believes that the 80/20 Rule is chronically misunderstood, for three reasons. First, it's almost never exactly 80/20. Most of the large-inventory markets are 80/10 or even less (no more than 10 percent of products account for 80 percent of sales). The second confusing fact is that the 80/10 doesn't add up to 100. The 80 and the 10 are percentages of different things, and therefore they don't need to equal 100. One is a percentage of products, the other a percentage of sales. Third, there is no standard convention on how to express the relationship between the two, or which variable to hold constant. Saying a market has an

---

[14] Wikipedia, the free encyclopedia. *Pareto distribution*. http://en.wikipedia.org/wiki/Pareto_distribution Extracted 4 May 2009.

[15] Martin, Roger L. (2007) *The Opposable Mind – How Successful Leaders Win Through Integrative Thinking*. Harvard Business School Press. Boston, Massachusetts. Pp. 76-77.

80/10 shape (10 percent of products account for 80 percent of sales) can be the same as saying it is 95/20 (20 percent of the products account for 95 percent of sales).[16]

The Long Tail – the alternative to the 80/20 Rule

The traditional interpretation of the 80/20 Rule is to only carry 20 percent of goods that account for the most sales. Chris Anderson's concept of the '**Long Tail'** is a reinterpretation of the 80/20 Rule. (Refer Figure 5). What the Long Tail offers, is the encouragement not to be dominated by the 80/20 Rule. In Long Tail markets, where the carrying costs of inventory are low, the incentive is there to carry everything, regardless of the volume of its sales. With good search engines the bottom 80 percent product can turn into a 20 percent product. Historically the traditional bricks and mortar retailer needed to focus on the 20 percent of products that sold well. That paradigm changed with computerisation and the Internet, as we in the west now live in the age of abundance, not in the age of scarcity. Distribution is now cheap.[16]

Some examples:
- Some of the most successful Internet businesses have leveraged the Long Tail as part of their businesses. For example, physical goods (e.g. Amazon, eBay), digital goods (e.g. iTunes, iFilm), advertising (e.g. Google), information (e.g. Google, Wikipedia), communities/ user-created content (e.g. MySpace).[17]
- Wikipedia is a Long Tail provider of information with 1 million articles in English compared with Britannica's 80,000. This means that a lot more specialist knowledge for different niches is now available.[17]
- Long Tails have proved a very popular concept, with micro-finance schemes (e.g. Bangladesh), user-driven innovation, economic modelling and marketing (viral marketing).[18]
- Rhapsody offers online music for 1.5 million tracks. Most music shops sell only 55,000 tracks at best.[17]
- Amazon offers 3.7 million book titles. A typical bookshop like Borders holds only 100,000 books.[17]

---

[16] Anderson, Chris (2006). The Long Tail – *Why the Future of Business is Selling Less of More.* Hyperion. New York. U.S.A. Pp. 120-132.
[17] Ibid., Pp. 23-89.

[18] Ibid., Pp. 229.

Figure 5

The Radical New Shape of Culture and Commerce:
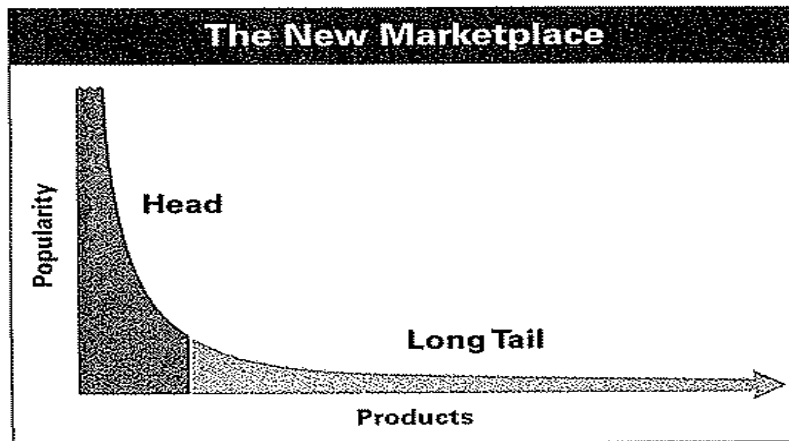
The Long Tail - The radical new shape of Culture and Commerce. (Anderson – 2006).

The Long Tail market (Refer Figure 6) provided by Amazon and Rhapsody provides them with 20 to 40 percent of their total sales which are products not available in offline retail stores. This is the fastest growing part of their businesses. There are lots of niche markets that have a steady demand.[17]

Figure 6
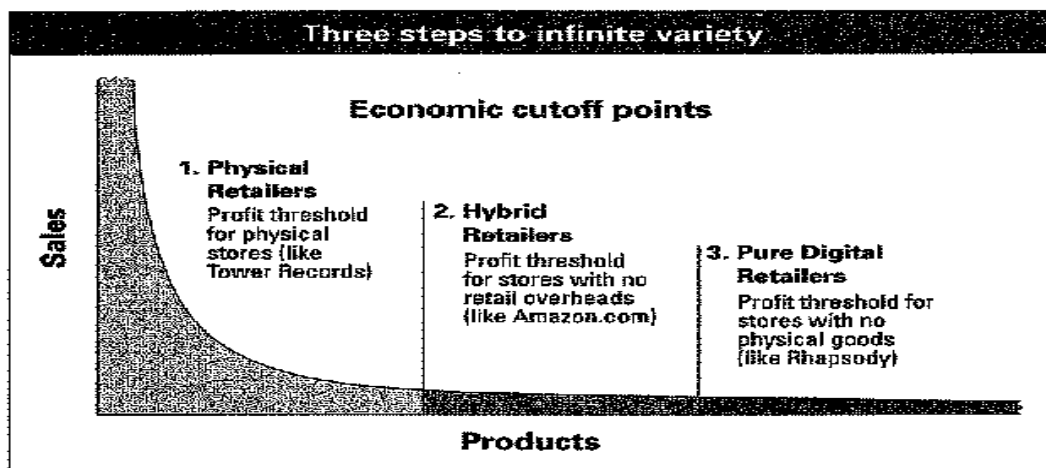
The Long Tail – Three steps to infinite variety. (Anderson – 2006).

Book sales show a typical long tail and power law distribution. For example, in 2004, 950,000 books out of the 1.2 million tracked by Nielson BookScan sold fewer than ninety-nine copies. Another 200,000 sold fewer than 1,000 copies. Only 25,000 sold more than 5,000 copies. The average book in America sells about 500 copies. In other words, about 98 percent of books are uncommercial, whether they were intended that way or not. Or expressed in a different way, each year only 20,000 of the 200,000 books published in English will make it into the average book store.[18] (Refer Table 3).

Table 3 - Book sales - Source: Book Industry Study Group[16]

Book Sales, 2004		
Sales Range	Titles	Units (millions)
1,000,000 or more	10	17.4
500,000 to 999,999	22	13.8
250,000 to 499,999	64	22.3
100,000 to 249,999	324	46.9
50,000 to 99,999	767	51.9
5,000 to 49,999	23,047	280.0
1,000 to 4,999	67,008	149.1
100 to 999	202,938	69.5
Sold 99 or less	948,005	14.3
Total	**1.2 Million**	**665 Million +**

The reality is that the vast majority of content (from music, books or movies) is not going to be hits, but there is still a demand for them, measured in hundreds of orders rather than millions. Sales follow a power law (they follow a 1/x log-log curve) – where a few products sell most and most sell a little. For example, only 10 titles (out of 1.2 million) in 2004 sold more than a million copies, whereas 1.15 million titles sold less than a 1,000 books each.[16]

In reality the power law curve isn't a straight line when people apply the 80/20 Rule. Box office revenues fall dramatically after rank 100 until they approach zero around 500. (Refer Figure 7). Movies don't get worse at rank 100 and the movie makers don't stop making movies at 500. An estimated 13,000 feature films are made each year in the U.S. alone. What happens is that the films past the first 100 or so simply failed to get much theatrical distribution. Or, to put it another way, the 'carrying capacity' of the U.S. theatrical industry is only about 100 films a year.[16]

Figure 7 - The Long Tail – Hollywood box office, 2005. (Anderson – 2006).
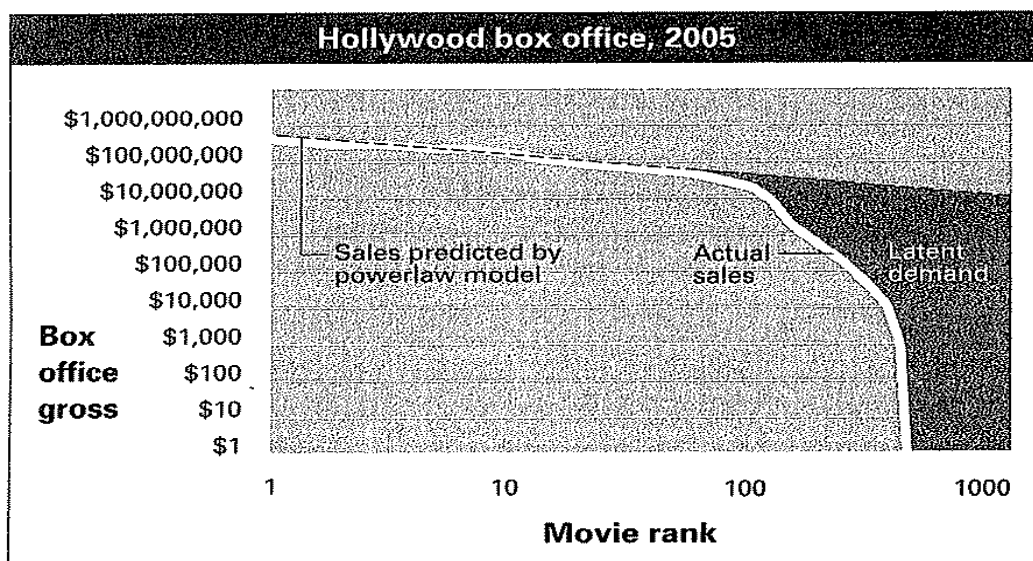
Chapter 6 - Zipf's law – Frequency of words used and size of cities

Zipf's law – Power Law of the frequency of Words

Zipf's law states that the frequency of any word is inversely proportional to its rank in the frequency table. Thus the most frequent word will occur approximately twice as often as the second most frequent word, which occurs twice as often as the fourth most frequent word. For example, in the 'Brown Corpus', which is a large database of over one million present day American English words, 'the' is the most frequently occurring word (7%), second is 'of' (3.5%) , followed by 'and' (2.9%). Only 135 words are needed to account for half the Brown Corpus. Zipf's law applies to many types of data in the physical and social sciences. It is one of a family of related discrete power law probability distributions. (Refer Figures 8 and 9 for log-log power law graphs). The simplest case of Zipf's law is a 1/f function, where frequencies are sorted by frequency from most common to least common. The second most common frequency will occur ½ as often as the first, the third will occur 1/3 as often as the first.[1]

Figure 8
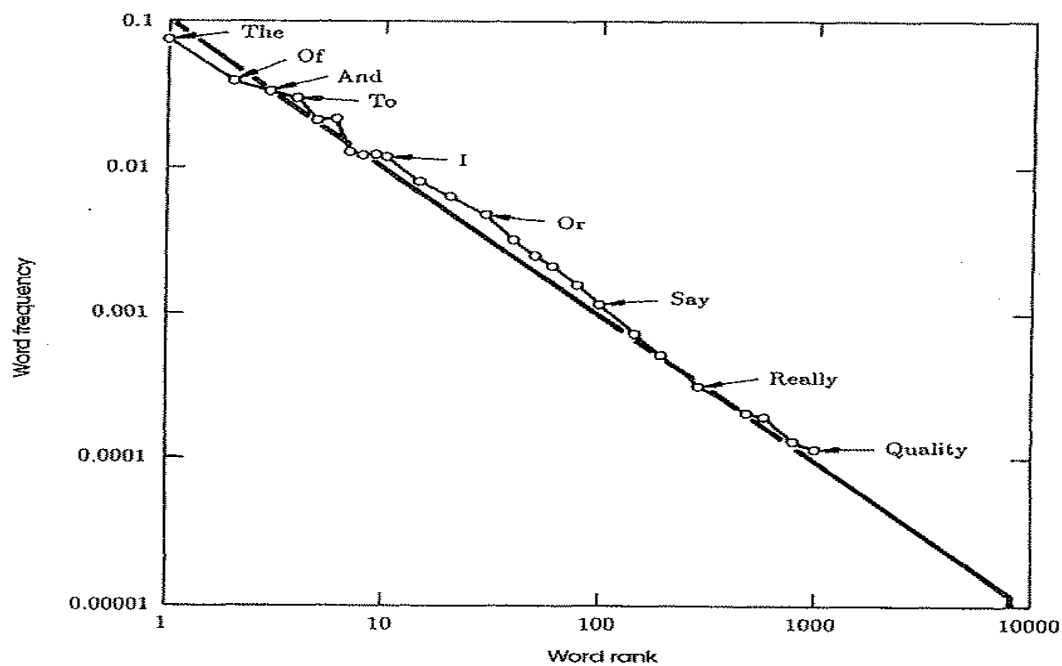
Power law curve for language – Zipf's law for the English language. (Casti – 1997).

---

[1] Wikipedia, the free encyclopedia. (2009) *Zipf's law*. http://en.wikipedia.org/wiki/Zipf%27s_law

49

Figure 9
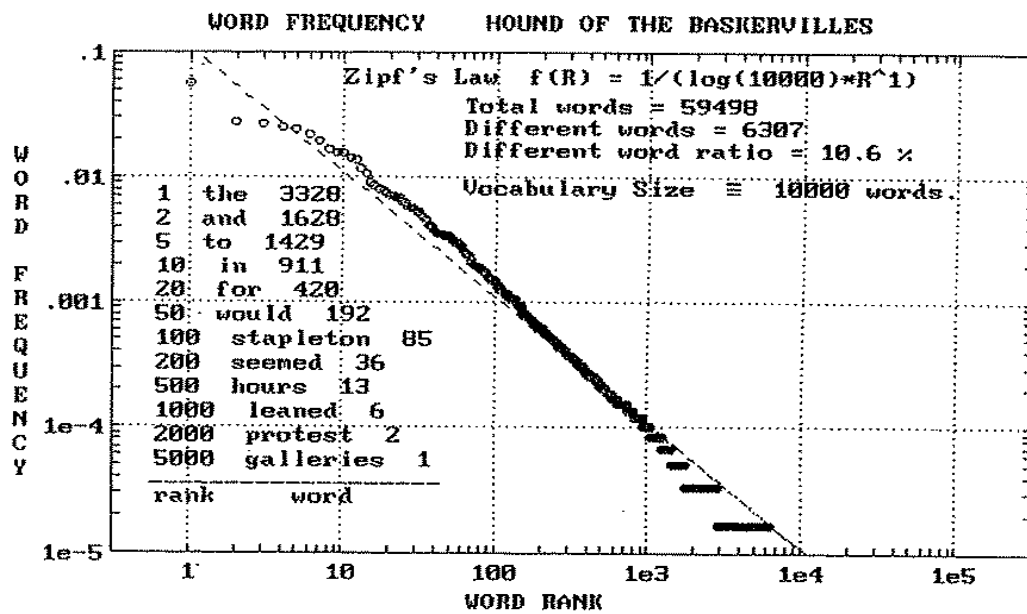
Power law curve for language – Zipf's law and the 'The Hound of the Baskervilles'. (Casti –1997).

Explanations for Zipf's law of words used

George Zipf (who the law is named after) was a Harvard linguist and suggested in the 1940s that the more you use a word, the less effort you find to use that word again, so you borrow words from your private dictionary in proportion to their past use. This would explain why out of the sixty thousand main words in English, only a few hundred constitute the bulk of what is used in writing, and even fewer appear regularly in conversation.[2]

Zipf's law is an empirical relationship which shows an emergent pattern that is not obvious from looking at the individual words of a language itself.[3]

Zipf's law, also called the rank size rule, is a power law that is related to the distribution known as the Fibonacci sequence and to that of the equiangular spiral. In terms of the Fibonacci sequence, each term is approximately 1.618 (the Golden ratio) times the preceding term. A special case of the Fibonacci sequence is the Lucas sequence consisting of these sequentially additive numbers 1, 3, 4, 7, 11, 18, 29, 47, 76, 123, 199 etc. When any log-linear factor is ranked, the ranks follow the Lucas sequence and each of the terms in the sequence can also be approximated by the successive values of powers of 1.618. For example, the third term in the sequence above, 4 is approximately 1.618^3 or 4.236 (which is approximately 4); the fourth term in the sequence, 7, is approximately 1.618^4 or 6.854 (which is approximately 7); the eight term in the series, 47, is approximately 1.618^8 or 46.979 (which is

---

[2] Taleb, Nassim Nicholas (2007). *The Black Swan – The Impact of the Highly Improbable.* Penguin Books. London. England. Pp. 219.

[3] Casti, John L (1997). *Would-Be Worlds. How Simulation is Changing the frontiers of Science.* John Wiley and Sons, Inc. New York. U.S.A. Pp. 128.

approximately 47). With higher and higher values, the figures converge. The rank size rule 'works' because it is a 'shadow' or coincidental measure of the true phenomena. The true value of rank size is thus not as an accurate mathematical measure but as a handy measure or 'rule of thumb' to spot power laws.[4]

Wide application of Zipf's law

Zipf found that the Zipf's law applies to a host of other phenomena, from population statistics to industrial processes. He analyzed Philadelphia marriage licences within one twenty-block area and showed that 70 percent of the marriages were between people who lived no more than 30 percent of the twenty-block distance away from each other. Other researchers have extended the rule to everything from atoms in a plasma, electricity bills, street addresses, stock prices, to the size of cities. At the heart of these observations is the ubiquity of power law distributions, the 1/f shape that Pareto first saw in his wealth curves.[5]

An interesting application of Zipf's law of words is its apparent use in determining if the 'noise' received from outer space is just 'noise' or if it is communication from another civilization. If the different sounds don't follow a Zipf distribution then the information is not considered a language and therefore not a message.

Zipf's law – Power law of size of Cities

If one ranks the population size of cities in a given country or in the entire world and calculates the natural logarithm of the rank and of the city population, the resulting graph will show a remarkable log-linear pattern. (Refer Figure 11). This is the rank-size distribution. For example a rank 3 city would have 1/3 the population of a country's largest city, a rank four city would have ¼ the population of the largest city, and so on.[4] The log (size) versus log (rank) will have a straight line slope of -1.[6]

---

[4] Wikipedia, the free encyclopedia (2009). *Rank-size distribution*. http://en.wikipedia.org/wiki/Rank-size_Law

[5] Anderson, Chris (2006). The Long Tail – *Why the Future of Business is Selling Less of More.* Hyperion. New York. U.S.A. Pp. 126.

[6] Axtell, Robert and Florida, Richard (March 2006). *Emergent Cities: Micro-Foundations of Zipf's Law.* Brooking Institution. U.S.A. Pp. 1-12.

Figure 11
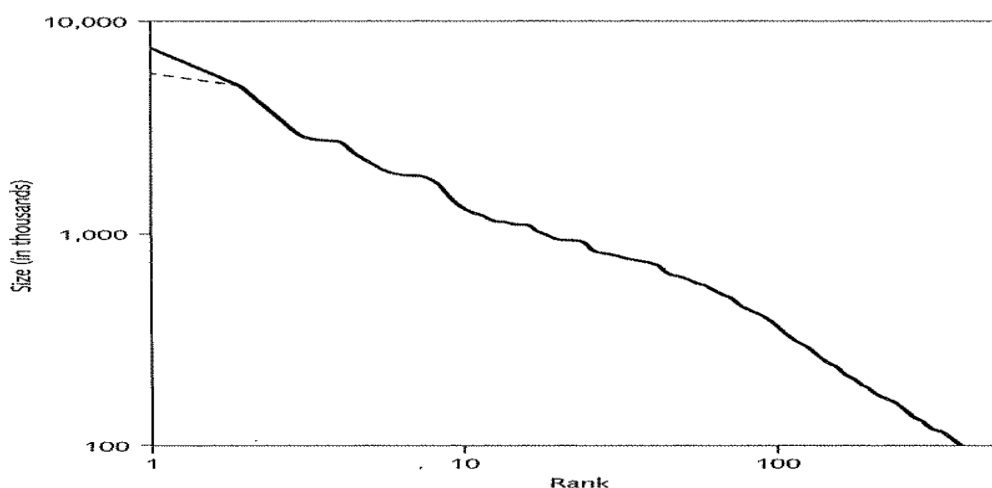
Power law curve of city size – Ranking of cities by size around the year 1920, from Zipf. The curve shows the number of cities in which the population exceeds a given size or equivalently, the relative ranking of cities versus their population. (Bak – 1997).

For example, there is a striking empirical regularity that governs city sizes in the United States. Cities are defined as metropolitan areas, rather than cities as political units. Figure 10 plots the log of each metropolitan area's population against the log of its rank (i.e. New York is 1, Los Angeles 2, and so on) for the 130 largest such areas in the United States in 1991. The relationship is impressively linear, and very close to a 45- degree line. If you multiply the rank of each city by its population, you get a very nearly constant number. For example, the 10th largest metropolitan area in the United States is Houston, population 3.85 million; the 100th is Spokane, population 370,000.[7]

---

[7] Krugman, Paul (1997) *How the Economy Organizes itself in Space: A Survey of the New Economic Geography* – Part of *The Economy as an Evolving Complex System II.* (Editors Arthur, W. Brian; Durlauf, Steven N.; Lane, David A) - Proceedings Volume XXVII. Santa Fe Institute. Studies in the Sciences of Complexity. Addison-Wesley. The Advanced Book Program. Reading Massachusetts. Pp. 239-262.

Figure 10
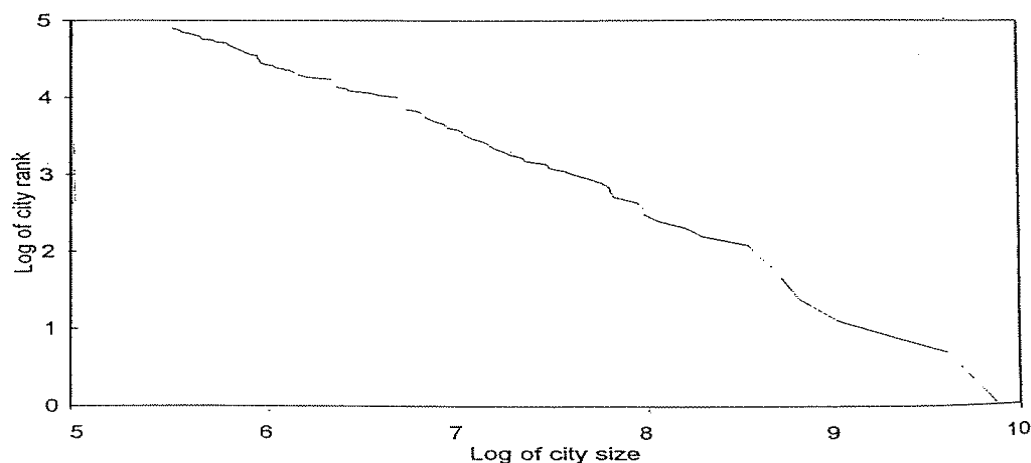
Power law curve for the size distribution of U.S. metropolitan areas from Krugman. (Arthur – 1997).

This statistical regularity has shown to be valid for most industrial countries today, as well as for individual countries over extended periods of time. There are well known exceptions to the law (as evidenced by the dotted line at the top left of Figure 11); often the largest city in a country is 'too large' to fit well into the overall size distribution and some countries have size distributions that deviate systematically from the Zipf distribution.[6] For example, the rank-size rule works reasonably well for the 31 largest urban areas in New Zealand, though the Auckland region is slightly larger than it predicts.[8]

Whilst noting the exceptions, it is clear that the Zipf law for city size and the Zipf law for firm size are two famously skewed size distributions and are among the single most regular statistical features of modern societies.[6]

Explanations for Zipf's law of city sizes

Axtell and Florida believe that despite considerable research, there does not exist a traditional explanation for Zipf's law based on behaviourally-credible micro-foundations. Whilst economists have recognized this uneven, hierarchal structure, they have been unable to generate empirically-salient explanations for it, based on the linear assumptions of rationality and equilibrium which are the standard economic assumptions.[6]

The purely neoclassical theory tells us that high-tech firms will tend to distribute themselves evenly across the landscape. There's no reason for any of them to prefer one location over another. But in real life, of course, they flock to places like California's Silicon Valley and

---

[8]   Small, John and Sweetman, Chris (September 2008). *Drivers of Economic Growth in Auckland. A Report for the Royal Commission on Auckland Governance.* Covec Limited.

Boston's Route 128 to be near other high-tech firms. This is due to positive and negative feedback loops of non-linear systems.[9]

Complexity explanation of Zipf's law

Like most agglomerations, Silicon Valley owes its existence to small historical accidents that, occurring at the right time, set in motion a cumulative process of self-reinforcing growth. Krugman believes the spatial economy is, self-evidently, a self-organizing system characterized by path dependence. It is a domain in which the interaction of individual decisions produces unexpected emergent behaviour at the aggregate level. Krugman believes the spatial economy is a dynamic landscape that is typically rugged with an evolution that involves 'punctuated equilibria', in which gradual changes in the driving variables leads to occasional discontinuous change in the resulting behaviour.[7]

Cities grow, according to Waldrop, because they build up slowly as small chance events become magnified by positive feedbacks which are non-linear random processes.[9]

Buchanan is clear as to the implication of Zipf's law. The implications of this power law pattern are as usual: there is no 'typical' size for a city in the USA or elsewhere, and no reason to see special historical or geographical situations behind the emergence of the very biggest. The growth of a city is a critical process, poised on the edge of great instability. There was probably nothing inevitable or special about the beginnings of New York, Mexico City or Tokyo. If you rewind history and play it again there would no doubt be great cities, but in different locations and with different other names. Even so, the power law pattern of cities would remain the same.[10]

Likewise, the more people aggregate in a particular city, the more likely a stranger will pick that city as his destination. The big get bigger and the small stay small, or get relatively smaller.[2]

The Axtell/ Florida complexity model that fits the Zipf data shows how cities and firms form and co-evolve from a simple micro-foundation: people form firms and these firms agglomerate to make cities. Agents have heterogeneous abilities; exhibit bounded rationality and interact directly with one another in out of equilibrium or non-linear environments.[6]

---

[9] Waldrop, M. Mitchell (1992). *Complexity – The Emerging Science of Order and Chaos.* Simon and Schuster. New York. U.S.A. Pp. 36-45.

[10] Buchanan, Mark (2000). *Ubiquity – The Science of history…Or Why the World is Simpler Than We Think.* Weidenfeld and Nicholson. London. England. Pp. 158.

Chapter 7 - Zipf's law – size of Companies

Robert Axtell used as data the entire population of tax paying firms in the U.S., being 5 million firms and 105 million employees. (Refer Figure 12). He found that the Zipf distribution characterizes firm sizes, with the probability a firm is larger than size 's', is inversely proportionate to 's'. Axtell found that the power law exponent 'α' was close to 1. These results hold for data from multiple years and for various definitions of firm size.[1]

Figure 12
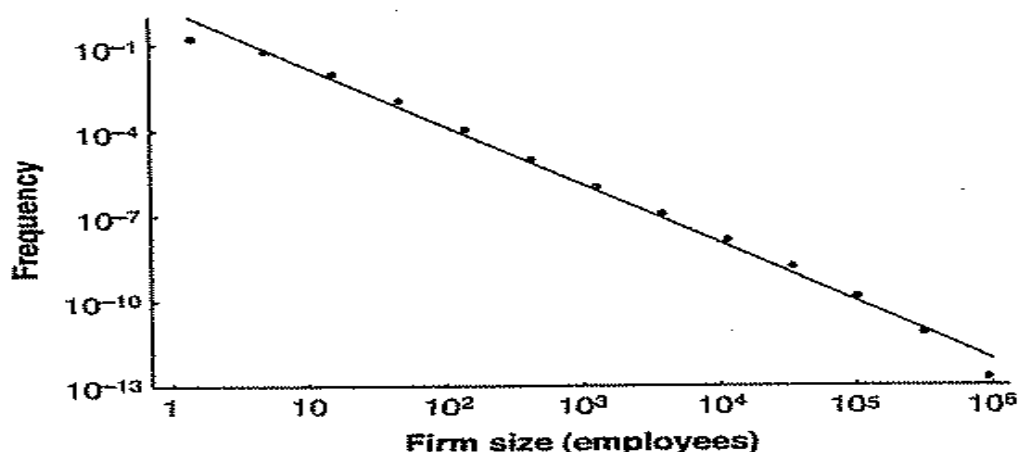
Power law curve for U.S. Company size distribution by employees (probability mass function). Data from 1997 U.S. Census. (Axtell – 2001).

Fitting the Zipf distribution means that firm size is highly skewed. The upper tail of the firm size distribution has often been described by the Yule or Pareto (also known as power laws or scaling) distributions. Whilst the Figure 12 graph shows a slight imperfection with too few very small and very large firms with respect to the Zipf fit, presumably due to finite size effects, it is a strong power law distribution.[1] When fitting data to scaling laws it is common that data at the extremes do not fit well due to so-called finite size cut-offs.[2]

Company size means either sales turnover or employee numbers

Axtell found that it does not matter whether one uses numbers of employees or the companies sales volume in constructing company size distribution functions, as the power law exponent of the distributions are approximately the same.[2] (Refer Figures 12 and 13).

---

[1]   Axtell, Robert L. (2001). *Zipf Distribution of U.S. Firm Sizes*. Volume 293 Science, 7 September 2001. www.sciencemag.org Pp. 1818 – 1820.

[2]   Axtell, Robert (June 1999). *The Emergence of Firms in a Population of Agents: Local Increasing Returns, Unstable Nash Equilibria, and Power Law Size Distributions.* Centre on Social and Economic Dynamics, Working Paper No. 3. June 1999. Brookings Institution. U.S.A. Pp. 40-44.

The relationship between firm output and size by number of employees once plotted on a log-log graph show that the two measures are highly correlated. There is significant variability at large sizes since these data bins hold few observations. The graph highlights the hypothesis of constant returns.[2]

Figure 13
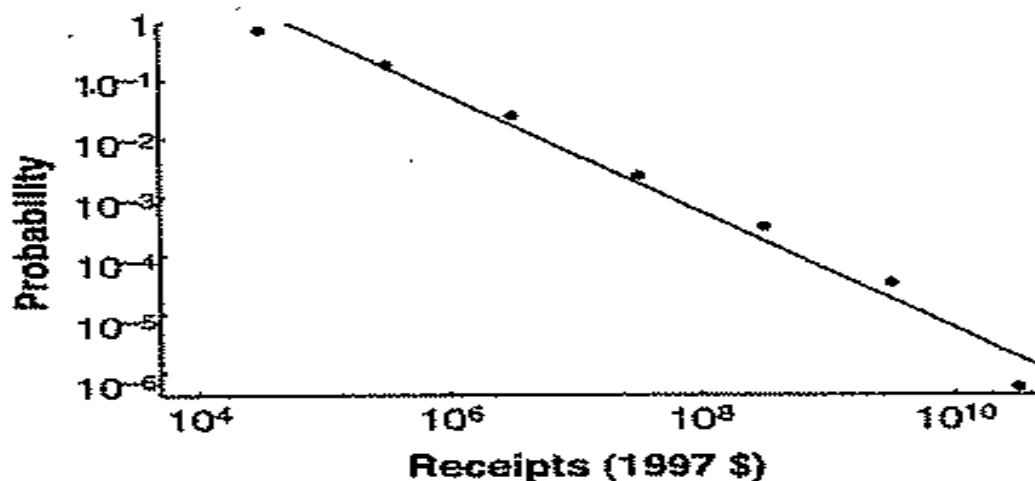
Power law curve for U.S. Company size distribution by revenue (probability mass function). Data from 1997 U.S. Census. (Axtell – 2001).

'Equilibrium' assumptions fail to explain Zipf's law of firm size

According to Axtell, there does not exist today a purely microeconomic explanation for the overall firm size distribution. The neoclassical theory of the firm distils the multi-agent character of real firms down to a single rational actor, who faces completely specified technological options, and who acts to maximize profit by choosing inputs that minimize cost. There are many problems with this firm-as-production-function picture, from its static technology to its hyper-rationality of requiring perfectly informed and rational managers. None of this explains Zipf law of company size.[3]

Given the power law character of actual firm size distributions, it would seem that equilibrium theories of the firm, such as principal-agent models, will never be able to grasp this essential empirical regularity. Intrinsic to this theory is that the system is far from equilibrium, a necessary condition in order to observe non-Gaussian fluctuations. Power law distributions and Per Bak's highly idealized sandpiles provide the general idea that power law size distributions result from 'self organized critical' processes, in which a system arranges itself such that large fluctuations are always possible, and these occur not infrequently.[2]

Axtell studied the out-of-equilibrium dynamics of an agent-based computational model.

---

[3] Ibid., Pp. ii-6.

Axtell assumes agents glean data concerning their economic environment from members of their social networks, that is, through local interactions. This relatively new methodology facilitates modelling agent heterogeneity, non-equilibrium dynamics, local interactions (network/ spatial processes) and bounded rational behaviour.[3]

It is Axtell's claim that the preoccupation with equilibrium notions is largely responsible for the relative neglect of the overall size distribution of firms in industrial organization. While the existence and stability of this distribution have been well-known for decades, and while microeconomic conceptions now dominate industrial organizational theory, there apparently does not exist a microeconomic (equilibrium) explanation of the aggregate size distribution. Indeed, perhaps it is the case that no equilibrium theory could ever reproduce the empirical data. Power law effects come from non-equilibrium distributions.[4]

Real firms have provided the important concept of bounded rationality, a notion still far from fully-integrated into economic theory.[4]

The new non-equilibrium models assume that agents can be arbitrarily heterogeneous. They can interact with one another directly (local interactions) or indirectly through aggregate economic variables (global interactions). Agents possess only a limited amount of information and are of necessity boundedly rational, since to model full rationality is too computationally complex in all but extremely simple environments. Attainment of equilibrium is not required.[4]

No traditional theory of economics comes close to explaining these effects, which should be of no surprise, since traditional economic theories hold onto the idea of economics existing in a world that they believe is static and in 'equilibrium'.[5]

Traditionally, economists are inclined to look for Nash equilibria in their models – these, they think, determine behaviour in the real world. But Robert Axtell's model has no stable Nash equilibria. That is to say, it can never settle down into an unchanging state. There is a constant flux as firms boom and bust. This model is a non-equilibrium one, and this makes it quite different from most microeconomic models of firm growth. Nothing certain can be deduced about the system. Rather, it means that we are forced to draw statistical inferences. We can not say, for example, how long it will take for a firm encompassing 25 per cent of the total labour force to appear. But we can, at any point, determine the probability of such a firm existing. Axtell's study was compiled from Census and Compustat data involving 20 million US firms in 1997, being the 5 million tax paying firms and 15 million firms with no payroll/ employees. Axtell's model produces results which are striking, since no other microeconomic

---

[4] Ibid., Pp. 87-90.

[5] Buchanan, Mark (2007). *The Social Atom – Why the Rich get Richer, Cheaters Get Caught, and Your Neighbour Usually Looks Like You.* Bloomsbury. New York. U.S.A. Pp. 180-186.

theory of the firm has correctly forecast the power-law nature of this distribution. The power law distribution is a consequence of purposive behaviour on the part of the individual agents in the market. For example, we can increase the size of each agent's friendship circle, or we can boost the factors that promote increasing returns of scale etc, but these things only alter the slope of the straight-line, but the log-log line stays straight.[6]

Companies getting bigger is unavoidable, since in a network economy the hubs (companies) must get bigger as the network grows.[7]

Scaling

Axtell found a power law pattern for sales of $1/S^2$. This means that firms with sales of $1 million are four times as numerous as those with $2 million, which are in turn four times as numerous as those with $4 million.[5]

Scaling means firms of size 1,000 are four times less common than firms of size 500, while firms of size 10,000 are 100 times less common than firms of size 1,000. While large firms are certainly important, as a practical matter approximately 80% of the U.S. workforce is employed in firms size 500 or smaller.[2] An interesting property of firm size distributions is that the qualitative character of such distributions is independent of how size is defined.[1] From what I know of the profile of New Zealand companies, I am sure these principles also apply to New Zealand.

Perpetual company churn

Any acceptable theory of firms should also account naturally for the way firms continually came and go. Most firms are ephemeral. Although General Motors, Microsoft and Exxon seem like permanent features of the economic landscape, they can not be comfortable.[5]
Of the largest 5,000 US firms operating in 1982, for example, only 35 per cent still existed as independent entities in 1996. There is a high 'turnover' of companies, which many economic theories of the firm do not acknowledge.[6]

Whether labour markets or other markets, there is actually no real equilibrium, as markets will have many unpredictable fluctuations and crashes. A theory of the firm must necessarily be a non-equilibrium theory in which the future is never certain.[6]

---

[6] Ball, Philip (2004). *Critical Mass – How one thing leads to another.* Arrow Books. London. United Kingdom. Pp. 329-336.

[7] Barabasi, Albert-Laszlo (2002). Linked – The New Science of Networks. Perseus Publishing. Cambridge, Massachusetts, USA. Pp. 200.

For the emergent power law pattern, the details don't matter

Out of all the chaos of millions of people doing an infinite variety of different things, emerges a mathematical pattern.[5] Axtell notes that this skewness has been robust over time, as it has lasted over six decades of firm size (from 10^0 where there are no employees to 10^6 employees). The skewness appears to be insensitive to changes in political and regulatory environments, immune to waves of mergers and acquisitions, and unaffected by surges of new firm entry and bankruptcies. It even survived large-scale demographic transitions within the work forces (e.g. women entering the labour market) and widespread technological change. Individual firms migrate up and down the Zipf distribution, but economic forces seem to have rendered any systematic deviations from it short-lived. Whilst the overall emergent pattern remains constant over time, at the individual level of firms there is perpetual churn. The other key point to note is the fact that details don't matter.[1]

Prediction of specific events and details is impossible, but it is possible to make high level predictions from the statistics that emerge and hold true for many power law events, i.e. wealth inequality, size of firms etc.[5]

Power law patterns offer an aid in getting to grips with systems that simply never seem to settle down, that are always churning and changing and evolving, where nothing seems constant, and nothing is ever repeated. Virtually all of past and current social science is based on the notion of equilibrium, which poses some difficulties if much of the world is not in equilibrium, perpetually evolving as feedback drives the growth of patterns that replace the old, only to be replaced again in the future.[5]

Chapter 8 - Complexity Economics

Beinhocker suggests that rather than portraying the economy as a static equilibrium system, complexity models present the economy as a buzzing hive of dynamic activity, with no equilibrium in sight. Just as the pattern of a whirlpool arises from interacting water molecules, these models showed complex patterns of boom and bust and waves of innovation emerging from the interactions of simulated agents, just as they do in the real economy. Viewing the economy as a complex adaptive system provides us with a new set of tools, techniques, and theories for explaining economic phenomena.[1]

According to Beinhocker[2] 'complexity economics' has five 'big ideas'. (Refer Table 4). These five ideas can not be analyzed using mathematical theorems alone. Complexity economics uses a broad toolkit of approaches, plus big doses of computer power. In addition, complexity researchers have imported new mathematical and statistical tools from physics, biology, and other fields to help them better understand the economy as an open, dynamic system.[2]

Table 4: Beinhocker's Five 'Big Ideas' that distinguish Complexity Economics from Traditional Economics

	Complexity Economics	Traditional Economics
Dynamics	Open, dynamic, non-linear systems, far from equilibrium	Closed, static, linear systems in equilibrium
Agents	Modelled individually; use inductive rules of thumb to make decisions; have incomplete information; are subject to errors and biases; learn and adapt over time (or not, as the case may be)	Modelled collectively; use complex deductive calculations to make decisions; have complete information; make no errors and have no biases; have no need for learning or adaptation (are already perfect)
Networks	Explicitly model interactions between individual agents; networks of relationships change over time	Assumes agents interact only indirectly through market mechanisms e.g. auctions.
Emergence	No distinction between micro- and macroeconomics; macro patterns are emergent result of micro-level behaviours and interactions	Micro- and macroeconomics remain separate disciplines.
Evolution	The evolutionary process of differentiation, selection, and amplification provides the system with novelty and is responsible for its growth in order and complexity	No mechanism for endogenously creating novelty, or growth in order and complexity

---

[1] Beinhocker, Eric D. (2007). *The Origin of Wealth. Evolution, Complexity, and The Radical Remaking of Economics.* Random House Business Books. London, Great Britain. Pp. 19-40.

[2] Ibid., Pp. 96-106.

The fact that we can get such widely varying behaviours simply from tweaking one variable demonstrates an important characteristic of nonlinear dynamic systems; sensitivity to initial conditions. A related characteristic is that nonlinear dynamic systems are path dependent. These two characteristics, sensitivity to initial conditions and path dependence, make nonlinear dynamic systems notoriously difficult to work with and in many cases impossible to predict.[2]

Joseph Schumpeter's 'Gales of creative destruction'

Technological innovation has been found to be the major force in economic growth. Studies show that growth in inputs (i.e., in capital and labour), only accounts for about 15 per cent of the growth in the economy. The other 85 per cent was technological innovation, which has been the major force in growth of output in highly industrialised economies.[3]

The Neoclassical economists tended to view innovation as an external or exogenous factor: a random variable that affected the economy – like the weather – but was outside the bounds of economic study. Schumpeter, however, believed that innovation had to be viewed as internal, or endogenous, to the economy and central to its understanding. He insisted that for growth to occur, there must be a source of energy within the economic system which would of itself disrupt any equilibrium that might be attained. For Schumpeter, that source of energy was the figure of the entrepreneur, whom he wrote about in almost heroic terms.[1]

According to Schumpeter, technological progress occurred in a random stream of discoveries. The commercialisation of new technologies, however, faced numerous barriers, ranging from the need for financing to the intransigence of old habits and mind-sets. Thus, like water behind a dam, the random rain of discoveries built up over time. (These metaphors compare very closely with Thomas Kuhn's later paradigm shifts in science). In Schumpeter's theory, entrepreneurs played the role of dam breakers, unleashing a flood of innovation into the marketplace. In this way, growth comes to the economy not in a steady stream, but as Schumpeter famously put it, in 'gales of creative destruction'. The origin of wealth, according to Schumpeter, lies in the heroic efforts of individual entrepreneurs. Schumpeterian wealth creation occurs when people like Richard Arkwright, Henry Ford, Thomas Edison, and Steve Jobs battle the odds to turn the technologies of their time into successful commercial enterprises.[1]

The entrepreneur, Schumpeter wrote, is 'the pivot on which everything turns'. Entrepreneurs – whether they operate in big firms or small ones, old companies or start-ups – are the agents of innovation and creative destruction. Their projects are the wellsprings of new jobs, higher incomes, and general economic progress. But in releasing their creative energies, rising

---

[3] Rosenberg, Nathan (2004). *Innovation and Economic Growth.* O.E.C.D.

entrepreneurs shove older ones aside, destroying their dreams and often their fortunes. The majority of people, even in rich countries, will never become entrepreneurs. Some will not even find employment in successful enterprises. And sooner or later, most businesses will fail, sometimes damaging whole communities as well as individuals.[4]

In the rise of one set of entrepreneurs and concurrent decline of others, Schumpeter sees profound social implications. 'The upper strata of society are like hotels, which are indeed always full of people, but people who are forever changing.' Just as businesses go up and down, so too do entrepreneurs and their families.[5]

Schumpeter's ideas could not be translated into the rigorous language of mathematics and therefore have been unable to be 'modelled', so it could never be reconciled with the mathematical Neoclassical framework, which limited its impact to most economists. His work has been slow to be absorbed into the economic mainstream, but highly valued by those interested in entrepreneurship.[1]

Economic progress shows punctuated equilibrium in action

Schumpeter did not like the prevailing view of economics as working in a state of equilibrium. He thought that capitalist reality is first and last a process of change and had nothing to do with any concept of perfect equilibrium.[6] Schumpeter represents advances in economic reasoning as nonlinear. The world involved much more of discontinuity and seismic convulsions, rather than regularity and smooth progression.[7] Whilst most individual technological innovations have only an incremental impact, overall, technological innovation proceeds in patterns of calm and storm.[8]

Innovation continuously revolutionizes the economic structures from within, destroying the old order and creating a new one. This process of Creative Destruction is the essential fact about capitalism. Those revolutions are not strictly incessant; they occur in discrete rushes which are separated from each other by spans of comparative quiet. The process as a whole works incessantly however, in the sense that there always is either revolution or absorption of the results of revolution, both together forming what are known as business cycles.[6]

---

[4] McCraw, Thomas K. (2007). *Prophet of Innovation – Joseph Schumpeter and Creative Destruction.* The Belknap Press of Harvard University Press. Cambridge, Massachusetts, U.S.A. Front cover and Pp. 7.

[5] Ibid., Pp. 70-71.

[6] Schumpeter, Joseph A. (2009) *Can Capitalism Survive? Creative Destruction and the Future of the Global Economy.* First published 1942. First Harper Perennial Modern Thought. New York. U.S.A. Pp. 31-43

[7] McCraw, Thomas K. (2007), op. cit., Pp. 352, 461-466, 474.

[8] Beinhocker, Eric D. (2007), op. cit., Pp. 174-175.

Schumpeter believed that all companies react 'adaptively' to change, but creative 'responses' come only from innovative acts by entrepreneurs. Their innovations can take many forms: for example, new commodities, new forms of organisation, or new markets. Innovating firms do not arise evenly throughout the economy. Instead, groups of these firms emerge just after an organisational or technological breakthrough in a particular industry. Meanwhile, powerful elements in society resist major innovation, because they tend to wreak havoc on existing structures. As a result, 'the history of capitalism is studded with violent burst and catastrophes'. It is no gentle process of adjustment but something 'more like a series of explosions'. The building of a railway where none had existed, for example, 'upsets all conditions of location, all cost calculations, all production functions within its radius of influence'.[9]

Economic networks

All technologies depend on a web of relationships with other technologies. These interrelationships are not just technological, but economic. The economic web that has grown around the automobile, for example, includes industries ranging from steelmaking to oil, hotels, and fast food. Also, all technologies are inherently modular; a car, for example, is made up of an engine, a transmission, a body etc. Modules are then assembled into 'architectures'. Innovations in modules can enable new architectures (e.g. the microchip enabling the PC), but it is innovations in architectures (e.g. the PC itself) that tend to have the big catalyzing ripple effects on innovation. All these interlocking economic/ technology networks or webs are therefore subject to cascades of change leading to the emergent pattern of punctuated equilibrium.[8]

Strategy is about surviving the 'Red Queen Race'

Schumpeter said that, 'Strategy must be seen in its role in the perennial gale of creative destruction; it can not be understood irrespective of it or, in fact on the hypothesis that there is a perennial lull'. Schumpeter believed that a creative response as opposed to an adaptive response can never be predicted and is therefore indeterminate, shaping long-run outcomes in a country, industry, or firm. It often depends on the leadership of specific individuals and it changes social and economic situations forever. It creates new conditions that would never have been developed without it.[7]

The longer the industry remains relatively stable, the more convinced the management team becomes that it really understands what is going on and is able to predict where the business is going. The more convinced of this management becomes, the more precisely and specifically it tunes its assets, skills, and people to succeeding in the environment it has been experiencing. Then one day a new technology is developed, a competitor has an idea, or

---

[9] McCraw, Thomas K. (2007), op. cit., Pp. 254-255.

consumer tastes begin to shift – and no one in the senior management team even knows about it. The world starts to change. At first the pattern-recognizing minds of the company's leaders can't believe the changes are really occurring. They believe that the problems they are experiencing are just a temporary anomaly and demand more evidence- after all; they've had their industry figured out for a long time.[10]

Management suddenly finds itself in a new game and stuck with the wrong mental models, the wrong assets, and the wrong skills. The punctuated nature of change tricks our pattern-recognizing minds into thinking that the world is more stable than it really is.[10]

Beinhocker quotes a study that found that true competitive advantage is both rare and relatively short lived. That study found that only 5 percent of the companies in the sample ever achieved a period of superior performance lasting ten years or more. For example, there are as many books about why great companies succeed as why great companies fail. The same companies show up in both sets of books. The companies that survived were often companies that underperformed the overall S&P 500 index by 20%, so not great 'excellent' companies. This highlights the ephemeral nature of competitive advantage, and the incredible dynamism of markets as companies rise and fall.[10]

There is no such thing as a safe, stable industry in the long term. Schumpeter's ghost is alive and well and the 'gales of creative destruction' are blowing harder than ever. Competitive advantage is rare and short-lived in the biological world as well. In biological systems, species are locked in a never-ending co-evolutionary arms race (prey v predator) with each other. Biologists refer to such co-evolutionary spirals as Red Queen races, named after the Red Queen in Lewis Carroll's Through the Looking Glass. It was she who said, 'In this place it takes all the running you can do, to keep in the same place'. There is no such thing as winning a Red Queen race, the best you can ever do is run faster than the competition. In evolutionary systems, sustainable competitive advantage does not exist; there is only a never-ending race to create new sources of temporary advantage. Understanding this means we change our definition of an excellent company from one of continuous high performance for very long periods (an achievement that is almost non-existent) to one that can string together a series of temporary advantages over time – in other words, a strong runner in the Red Queen race.[10]

Noting there are virtually no excellent companies that last multiple decades and the extreme rarity of repeated excellence, brings us to a brutal truth about most companies. Markets are highly dynamic, but the vast majority of companies are not. Studies have found that change in the economy is driven more by the entry and exit of firms than by the adaptation of individual companies.[10]

---

[10]   Beinhocker, Eric D. (2007), op. cit., Pp. 328-334.

From the market's point of view, each business is an experiment in Business Plan space, some succeed and are amplified, while others fail and disappear. Markets tend to create more surprise and innovation than do corporations. Companies are 'Big Person' hierarchies with all their human foibles and distortions, while markets are almost pure evolutionary machines.[10]

Chance events are important

Beinhocker believes that chance events coupled with positive feedback rather than just technological superiority will determine economic superiority – not some abstrusely defined edge in a given area of expertise (i.e. the best technology or best product always wins) as required in classical economic models. Economic models which previously excluded randomness, now need to factor in, unexpected orders, chance meetings, and such things as management whims, since they all help determine which products achieve early sales and, over time, which firms dominate.[11]

---

[11] Taleb, Nassim Nicholas (2007). *Fooled by Randomness – The Hidden Role of Chance in Life and Markets. Second Edition.* Penguin Books. London. England. Pp. 176.

Chapter 9 - Share market fluctuations

Truncated power laws show up in stock price fluctuations

For the last fifty years, Finance theory (which I studied as an undergraduate) has stressed the importance of the 'Efficient Markets Hypothesis' and the random walk model. Nobel Prizes in Economics have been awarded to people who have contributed to this body of knowledge. Even in the last five years, we can see the continued importance of these Gaussian based theories in the global banking collapse which had relied on a sophisticated Gaussian based finance equation.

This reliance on Gaussian based models has been increasingly criticised as it does not completely account for reality. The part of reality that it does not account for happens to be very small but very dramatic, i.e. share crashes. A random walk model effectively never predicts such a crash because it is simply outside of the results it produces.[1]

Perhaps the resistance to move away from Gaussian based models is that share price fluctuations don't exactly follow a power law distribution either, but price movements fall between the two systems. (Refer Figure 14.) Finance theory therefore needs to accept that a modified or truncated power law applies to share price movements to account for share fluctuations ranging from tiny fluctuations to drastic events (such as market crashes). If shares prices followed a pure power law distribution then events like 'Black Monday' of 19 October 1987, when all firms lost from 20 to 50 percent of their market value, would be more frequent.

Figure 14
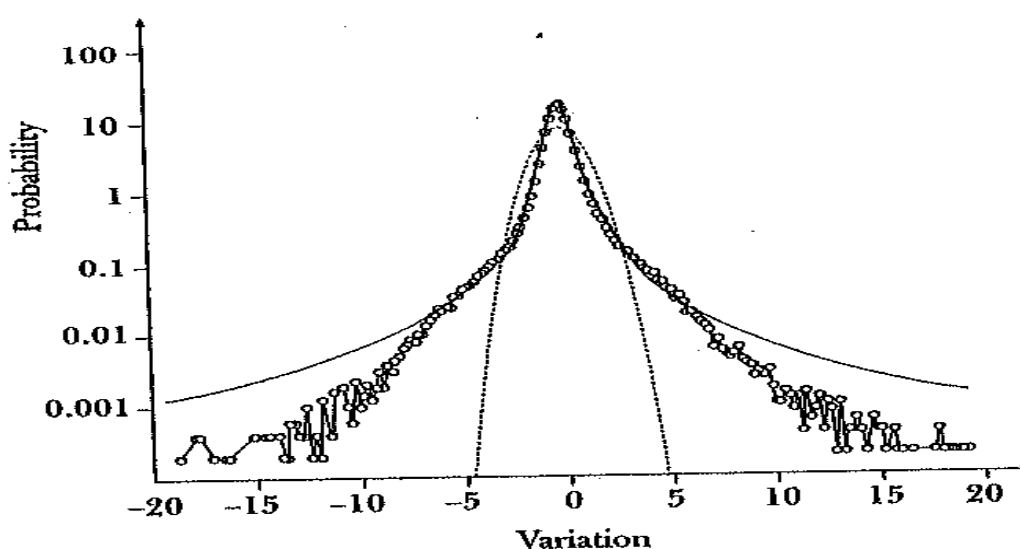
Shares – The probability distribution function for fluctuations in the S&P 500 markets index for a time interval of 10 minutes (same general shape as intervals of 1 minute and at least one day). For comparison, the probability distribution for Gaussian fluctuations (a random walk) is shown as a dotted line. The solid line shows the probability distribution for a power-law. (Ball – 2004).

Beinhocker cites Stanley et al. who used 40 million data points of every five minutes of stock transactions from 1994 to 1995 for one thousand of the largest U.S. companies. They found that the fluctuations in stock prices follow clear power laws in the tails of the distribution. Stanley varied his time periods over three orders of magnitude, from 5 minutes to 6,240 minutes (16 days). The data did look somewhat more Gaussian than the shorter periods, but it still very clearly followed a power law.[1] Just as the Gutenberg-Richter law (See Chapter 12) for the frequency of earthquakes shows that large earthquakes fall on the same power-law straight line as small earthquakes, so do share price fluctuations. You therefore do not need a separate theory for large earthquakes or large price fluctuations.[1]   (Refer Figure 15).

Figure 15
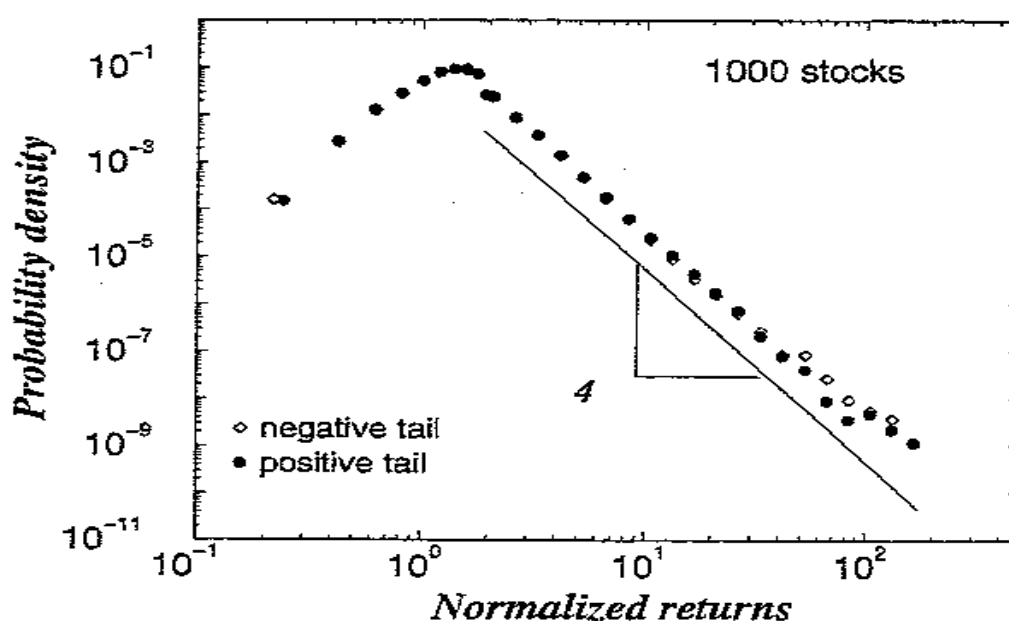
Power law curve for Shares – Probability density function of the normalized returns of the 1000 largest companies in the TAQ database for the two year period 1994-1995. (Stanley et al. – 2000).

Power laws show up in returns, volumes, numbers of trades and order flow

Whilst changes in share price fluctuations follow either truncated power law distributions or power law 'tails', other share market attributes also have pure power law distributions. For example, there is growing empirical evidence of the power-law tailed nature of distributions that characterize large movements of distinct variables describing stock market activity such as returns, volumes, numbers of trades, and order flow. These findings were established with 'dozens of millions of data points'.[2] Remarkably, the exponents that characterize these power laws are similar for different countries, for different types and sizes of markets, and for different market trends, suggesting that a generic theoretical basis may underlie these regularities. The Stanley team model links large movements in market activity to the power

---

[1]   Beinhocker, Eric D. (2007). *The Origin of Wealth. Evolution, Complexity, and The Radical Remaking of Economics.* Random House Business Books. London, Great Britain. Pp. 180-185.

law distribution of the size of large financial institutions. The trades made by large financial institutions create large fluctuations in volume and returns.[2]

The power law distribution of **trading volume** (shares traded) is 3/2 or 1.5, which is referred to as the '*half-cubic*' law of trading volume.[3] The Stanley team did empirical analysis on the 262 largest stocks listed in the London Stock Exchange (Refer Figure 16) and they confirmed that volume has a power law decay of 3/2 (the slope of Figure 16 being 2.4-1= 1.4) which is consistent with their results from the Paris Stock Exchange (Refer Figure 17).[3]

Figure 16
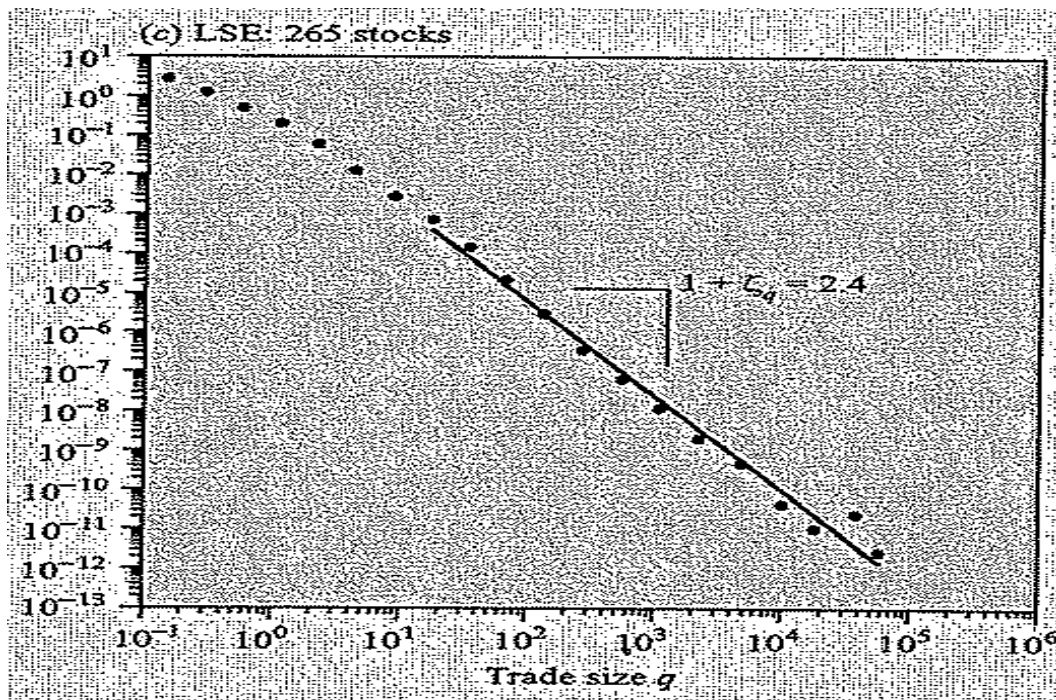
Power law curve for Shares – Probability density function of trade volumes for 265 stocks listed in the London Stock Exchange. (Plerou et al. – 2004).

---

[2] Gabaix, Xavier; Gopikrishnan, Parameswaran; Plerou, Vasiliki; Stanley, H. Eugene (2004). A *Theory of Large Fluctuations in Stock Market Activity*. February 17, 2004. Pp.1-35.

[3] Plerou, Vasiliki; Gopikrishnan, Parameswaran; Gabaix, Xavier; Stanley, H. Eugene (2004). *On the Origin of Power-Law Fluctuations in Stock Prices*. Published in Quantitative Finance 4 (February 2004) C11-C15. arXiv:cond-mat/0403067v1 (cond-mat.dis-nn) 2 Mar 2004.

Figure 17

(d) Paris Bourse

$1 + \zeta_q = 2.5$

Probability density $P(q)$
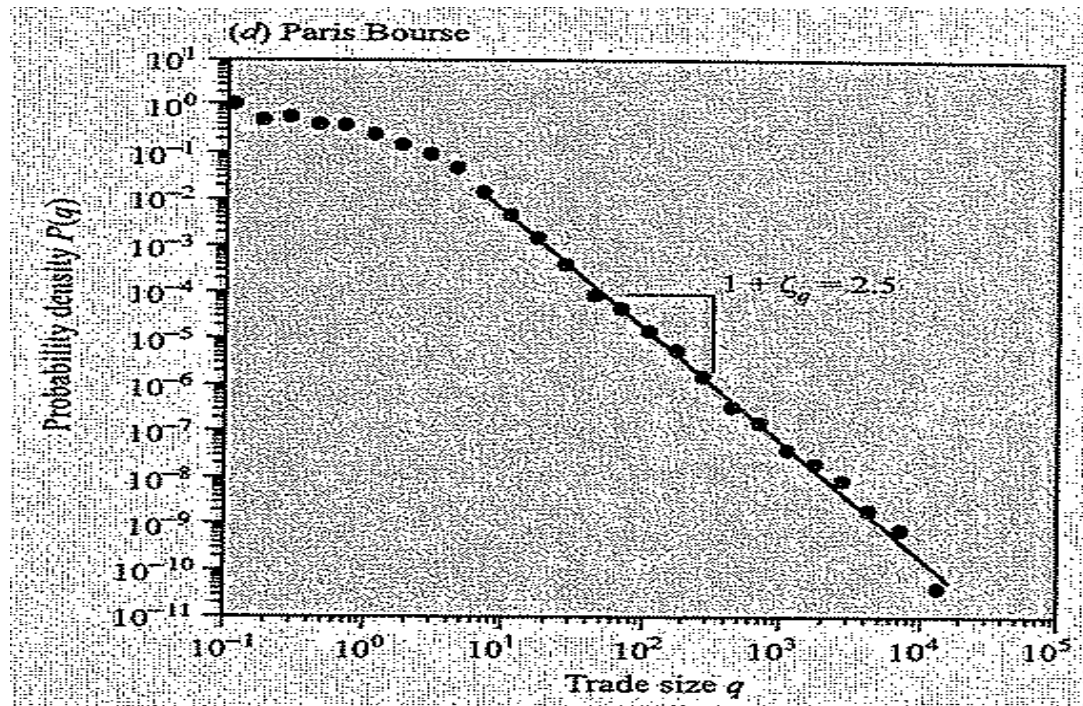
Trade size q

Power law curve for Shares - Probability density of the individual trade sizes q for 30 large stocks in the Paris Bourse from January 1995 to October 1999. This represents 35 million observations. (Gabaix et al. – 2004)

Figure 18

ζ

○ Nikkei '84–'97
■ Hang–Seng '80–'97
◇ S&P 500 '62–'96

Cumulative distribution

Normalized daily returns
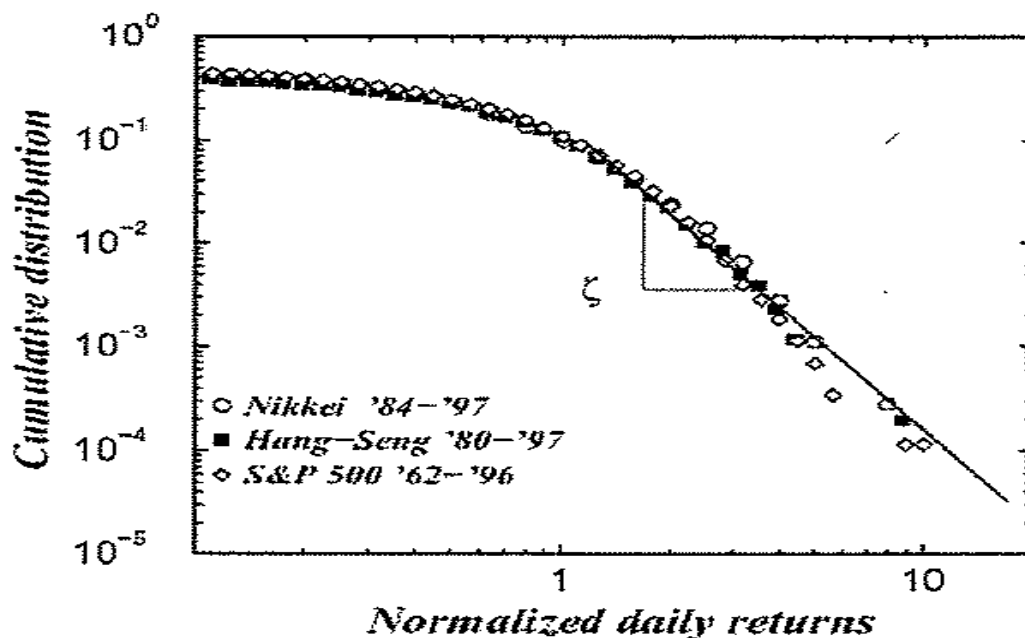
Power law curve for shares – Zipf plot for the daily fluctuations in the Nikkei (1984-97), the Hang-Seng (1980-97), and the S&P 500 (1962-96). (Gabaix et al. – 2004)

The power law distribution for **returns and the number of trades** is 3 (cubic law of returns). Refer Figure 18, which shows how the return distributions of three different countries are very

similar. The cubic law predicts the existence of a few very large returns over a century. The 1929 and 1987 stock market crashes are not outliers to the cubic law. Hence the theory of the cubic law of returns might be a theory of crashes.[3]

The standard Efficient Markets Hypothesis cannot explain these results

All these wild fluctuations fly in the face of the efficient market hypothesis, and the notion that markets are in equilibrium. In equilibrium, fluctuations should be small.[4] Gaussian, random walks almost never have fluctuations greater than five standard deviations, yet in real economic data, such as stock market crashes, five standard-deviation events, and even larger-deviation ones, do in fact occur.[2] For example, the stock market crashed more than 500 points on a single Monday in October 1987. A lot of blame went to computerized trading, but computers had been around for years. Is there any reason why the crash came on that particular Monday?[5]

One popular idea is that most extreme movements might simply be due to 'external shocks', destabilizing events such as the terrorist attacks of 9/11, or major corporate or government scandals. Unfortunately many large market movements seem to take place in the absence of any big event. In 1991, when a group of economists examined the fifty largest one-day price movements in the United States in the years since World War II, they found that many of them happened on days when there was no big news.[6]

Complex adaptive system models help our understanding

Economic systems are examples of complex interacting systems for which a huge amount of data exists.[1] Markets are complex systems for which the variables characterizing the state of the system – e.g. the price of the goods, the number of trades, and the number of agents, are easily quantified. These variables serve as fundamental examples of scale invariant behaviour – the scaling laws are valid for time scales from seconds to decades.[7]

The power law implies that there is no such thing as a typical fluctuation, and so there is no reason to think that every swing either up or down is in any sense unusual. Waldrop believes that the idea that sudden and terrifically large changes need explaining doesn't always hold up. In a market organised to the critical point, even the great stock market crashes are simply ordinary, expected events – although it is true that we should expect them infrequently.

[4] Buchanan, Mark (2000). *Ubiquity – The Science of history…Or Why the World is Simpler Than We Think.* Weidenfeld and Nicholson. London. England. Pp. 141-146.

[5] Waldrop, M. Mitchell (1992). *Complexity – The Emerging Science of Order and Chaos.* Simon and Schuster. New York. U.S.A. Pp. 9.

[6] Buchanan, Mark (2007). *The Social Atom – Why the Rich get Richer, Cheaters Get Caught, and Your Neighbour Usually Looks Like You.* Bloomsbury. New York. U.S.A. Pp. 70-79.

[7] Matia, Kaushik; Ashkenazy, Yosef; and Stanley, H. Eugene (2003). *Multifractal Properties of Price Fluctuations of Stocks and Commodities.* arXiv:cond-mat/0308012v1 (cond-mat.stat-mech) 1 Aug 2003. Pp.2.

Historically the big swings have not been easy to explain, except perhaps in hindsight. Even though there are no indications whatsoever, the market may fall by 20 per cent tomorrow. Such events need not be triggered by anything exceptional.[5] The cubic power law predicts the existence of large returns and stock market crashes such as the 1929 or 1987 crashes. They are not outliers.[3]

The major ups and downs in the market cannot always be traced to any external shocks or to the actions of any one person. Rather, the tendency for large fluctuations, a universal feature of all markets, emerges out of the way the agents, the atoms of this world, organise themselves into a delicate pattern of interdependence. Understanding why one rally is really big while another is small means getting down into the complex ecology of theories or hypotheses in the minds of all the traders, seeing how the actions of one trader, by influencing the prices, changed the behaviour of others, or did not.[7]

The real-world is a far more interesting place than the equilibrium world imagined by traditional economics. Oscillations that do not settle down, punctuated equilibrium, and power laws – these are all signature behaviours of a complex adaptive economy at work.[2] The economy is too complex, too nonlinear, too dynamic, and too sensitive to the twists and turns of chance to be amenable to prediction over anything but the very shortest of terms.[8] Since stock market fluctuations follow a power law, a small adjustment of interest rates or other change in background conditions could (if only rarely) be the trigger for a large swing in the market, or might have no effect at all.[9]

Commodity markets are similar to stock markets

Commodity markets have been found to have qualitative features similar to those of the stock market, despite having special features such as most commodities requiring storage and the need for transportation to markets etc.[8]

Per Bak discusses studies from the 1960s where the monthly variation of the cotton price is considered. Histograms were prepared, recording to how many months the variation would be 0.1%, how many months it would be 1%, how many months it would be 10%, etc. (Refer Figures 21 and 22). The result is a power law distribution (log-log chart) for the price fluctuations. The most important feature of this power law is that it copes with infrequent and very large events, just like the Gutenberg-Richter law for earthquakes (See Chapter 12). In Per Bak's view, the economists largely ignored this result, probably because they didn't have the faintest clue as to what was going on, since there is no possibility of having large

[8] Beinhocker, Eric D. (2007), op. cit., Pp. 323.
[9] Gribbin, John (2004). *Deep Simplicity. Chaos, Complexity and the Emergence of Life.* Allen Lane. An imprint of Penguin Books. London, England. Pp. 149-150.

fluctuations or catastrophes in an equilibrium economy with its inherent 'normal distribution' assumption.[10]

Figure 21
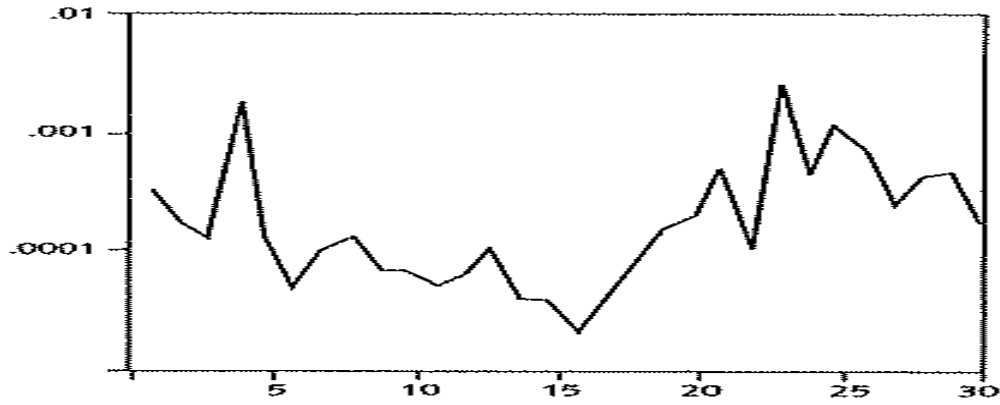
Monthly variations of cotton prices during a period of 30 months, from Mandelbrot. (Bak – 1997).

Figure 22
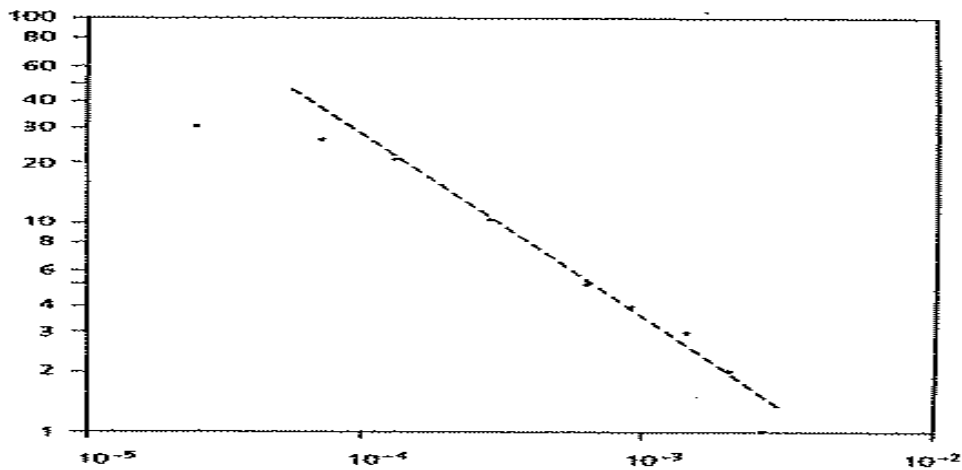
Power law curve for cotton price – The number of months where the relative variation exceeded a given fraction, from Mandelbrot. (Bak – 1997).

---

[10] Bak, Per (1997). *How Nature Works – The Science of Self-Organized Criticality*. Oxford University Press. Oxford, U.K. Pp. 186-192.

Chapter 10 - Company size fluctuations (Growth/ decline rates)

Quantifying business firm fluctuations

In 1996 physicist Gene Stanley and economist Michael Salinger looked at the growth rates of publicly traded US manufacturing companies between 1975 and 1991. Encompassing around 8,000 firms, this study found that growth rates followed a power-law relationship. This means that plotting the logarithm of the growth rate against the logarithm of the probability of that rate produces a straight line. Or in this case, two straight lines: one for positive growth rates and one for negative, both with the same slope. This generates a tent-shaped graph. They found that the power law holds true for two different measures of firm size – sales (revenue) and number of employees. Power laws are the universal signature of interdependence. A power law generally emerges from collective behaviour between entities through which local interactions can develop into a long-range influence of one entity on another.[1] (Refer Figure 23).

Figure 23
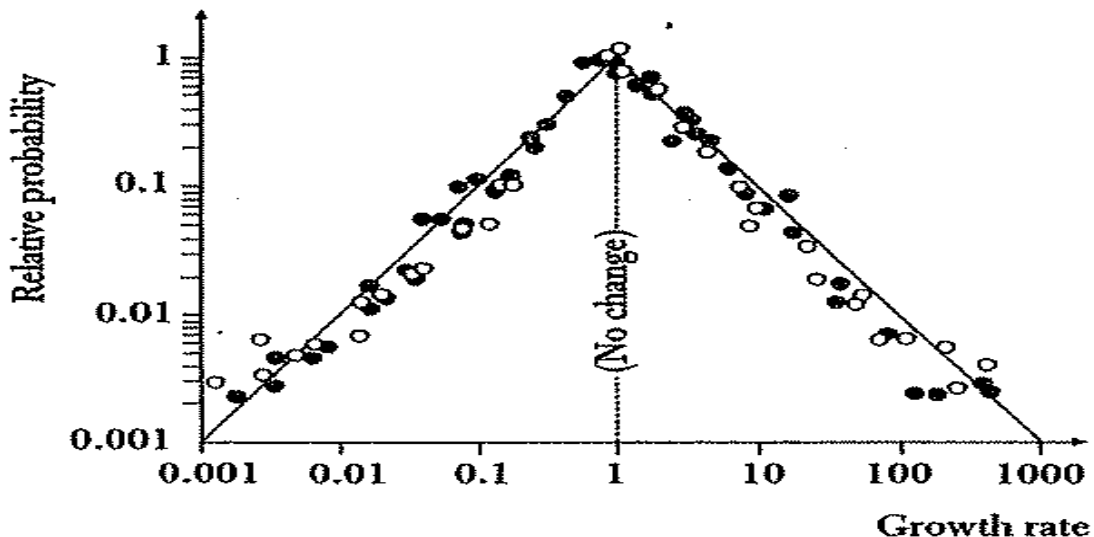
Tent shaped distribution of growth rates for all publicly traded U.S. companies between 1975 and 1991. The growth rates are measured by increases (or decreases) in both sales (black dots) and number of employees (white circles). The probability of a certain growth rate decreases as the rate increases according to a power law. Decreases in size follow the same power law as increases, making the graph tent shaped. From Stanley et al. (Ball – 2004).

Stanley and colleagues also analyzed the Computstat database comprising all publicly traded U.S. manufacturing companies within the years 1974-1993. They found the distribution of growth rates is different for different size firms, with a width that varies inversely with a power

---

[1] Ball, Philip (2004). *Critical Mass – How one thing leads to another*. Arrow Books. London. United Kingdom. Pp. 324-325.

of firm size. Similar variation is found for other complex organisations, including country size, university research budget size, and size of species of bird populations.[2]

Stanley and collaborators disregarded classical theory of firms which divided each business sector and they analyzed all firms together based on the 'complexity' viewpoint that all firms are actually interconnected. To find out growth patterns they made histograms that show how many times a firm grows by a given amount – e.g. the sales this year divided by the sales last year, sorted by the size of the firm. Large firms cannot grow or shrink very much in a given year. For example, General Motors could not grow by a factor of 10 in a single year. Big firms have narrow histograms and small firms can have a very wide histogram. The width, or standard deviation, of a histogram is a decreasing function of the size of the firm.[2]

Their analysis shows that the standard deviation is a function of the size of the firm, and once plotted against frequency, shows an approximately linear log-log graph result – suggesting that the deviation is a power law. The slope of the power law is on the order of 0.2. This number also seems to be quite universal, e.g. if you measure firm size by number of employees (rather than by annual sales) the slope is the same. It is therefore possible to collapse all these tent-shaped distributions onto the same scaling curve by using appropriate scaled variables.[2]

For some time it has been known that the 'economies' of many large business firms are the size of small countries, but it has been suggested that business firms might also have similar organization structures to countries. Different studies looked at economies of countries and found the identical tent-shaped distribution for the growing and shrinking of the economies of countries as business firms. So in this narrow sense, countries and business firms are identical. Funded scientific research in universities also grow and shrink in a tent-shaped distribution and when size-dependence is scaled out, so does the dynamics of North American breeding bird populations.[2] (Refer Figure 24).

[2] Stanley, H.E.; Amaral, L.A.N.; Gopikrishnan, P.; Plerou, V. (2000). *Scale invariance and universality of economic fluctuations. Boston University.* Physica A 283 (2000) Pp. 31-41. Elsevier Science B.V.

Figure 24
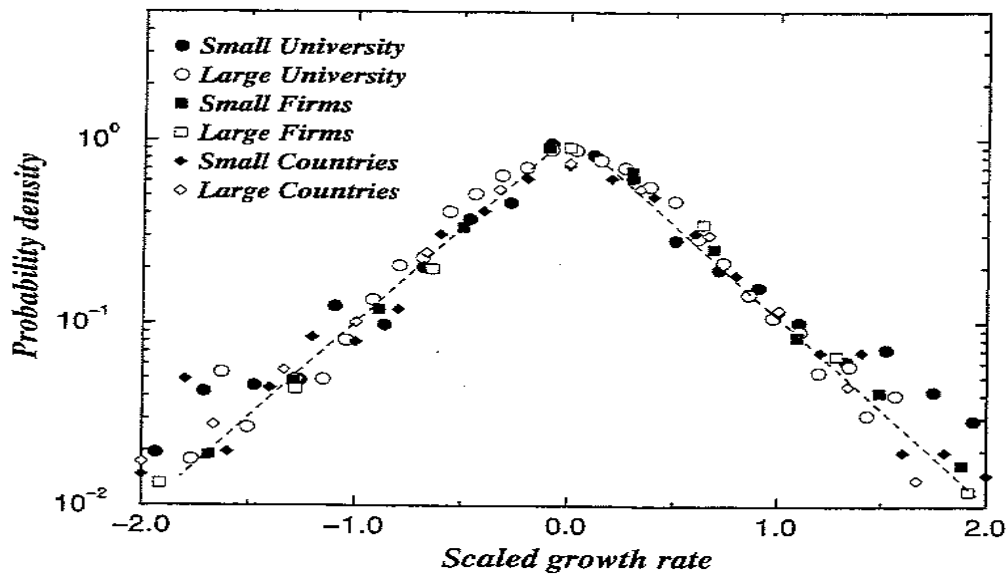
Tent shaped distribution of scaled annual growth for different organisations. (Stanley et al. – 2000).

The growth and shrinking of complex organisations appears to follow a universal power law that is immensely robust. This universal law demonstrates its validity for a wide range of organisation types – from business firms to bird populations. These results are great for modellers. Stanley believes that the empirically observed power law scaling does not require a critical state of the system, but may arise from an interplay between random multiplicative growth and the complex structure of the units composing the system.[2]

Growth rates of firms also follow another power law – they fluctuate less among larger firms than among smaller ones.[3]

---

[3] Buchanan, Mark (2007). *The Social Atom – Why the Rich get Richer, Cheaters Get Caught, and Your Neighbour Usually Looks Like You.* Bloomsbury. New York. U.S.A. Pp. 180-186.

Chapter 11 - Economic extinctions (Company failures)

Failure is the norm and is beneficial to the system

Failure is everywhere, across time, across place and across different aspects of life. 99.99% of all biological species that have ever existed are now extinct. Failure in this context is measured over hundreds of millions of years. On a dramatically shorter timescale, more than 10 per cent of all the companies in America disappear each year. Large and small, from corporate giants to the tiniest one-person business, they fail. However, for those that disappear there is a strong numerical bias of being at the newly formed and small end of the size and newness spectrum. Failure can be highly beneficial. Failure at the level of the individual component part can, paradoxically, enhance the fitness of the system as a whole.[1]

It is failure rather than success which is the distinguishing feature of corporate life. We see the survivors, and their triumphs are lionized. But the failures remain virtually forgotten. For example, Ormerod found that of the largest 100 US firms at the end of each decade from 1919 to 1979, only thirty-three out of the top 100 in 1919 remained in the list in 1979. The total sample size was only 216 companies, being the total number of US top 100 companies over the sixty-year period. Operating on a massive scale and possessing enormous resources, almost one in every four were unable to remain in the top 100 for more than a decade. Huge assets and years of successful operation, proved no guarantee, not merely of continued success, but of their very survival.[1]

The average lifespan of small firms is shorter than larger ones, but this is accounted for almost entirely by relatively high death rates in the very early years of existence. After the first few years of existence, the difference between large and small firms' survival potential narrows dramatically. Eventually most firms fail and the relationship that describes the pattern of extinctions amongst firms appears to be virtually identical to that which describes biological extinctions.[1]

Company failures show punctuated equilibrium and power law effects

Figure 24a (Data from Ormerod) shows company failures of the world's largest 100 companies and it shows very occasional large spikes, but mostly it has a background of a typically low level of extinction. It effectively shows a punctuated equilibrium pattern – mostly low extinction levels, punctuated at irregular intervals by much larger ones.[2] Spikes naturally seem to have occurred during prominent political or economic events such as World War 1,

---

[1] Ormerod, Paul (2005). *Why Most Things Fail – Evolution, Extinction and Economics*. Pantheon Books, New York, U.S.A. Pp. vii- 33.

[2] Ibid., Pp. 169-210.

recessions of the 1920's, the depression, World War Two, and oil crises of the 1970's and 1980's etc.

Figure 24a - Number of extinctions of the world's largest 100 companies in 1912, annual basis, 1912-95. (Ormerod 2005).
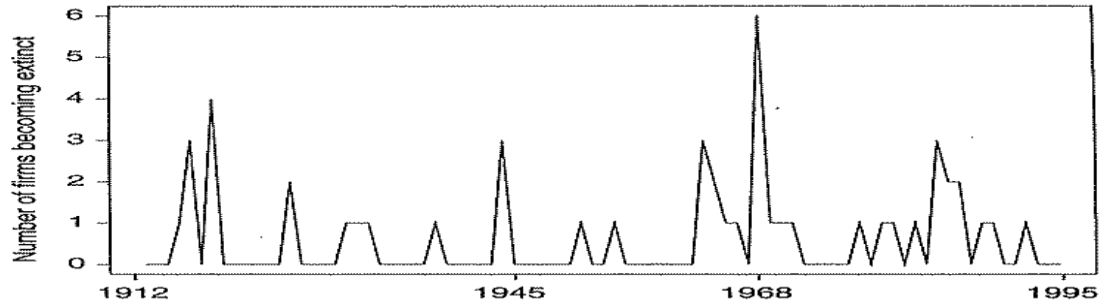
A standard way of presenting data organised by the frequency with which different values are observed is in a histogram. Figures 24b shows the histogram of the frequency of annual extinctions of the world's largest firms and Figure 24c plots the log of the rank of the size, so that the largest number is given rank 1, the second largest rank 2, and so on. This is the standard technique when the data series contains values of zero, which negates the possibility of using the logs of the raw data itself. The top left hand corner of the chart is labelled '0' which labels that point for the frequency of no extinctions for that year. '6' is the greatest frequency of extinctions in this sample in one year.[2] This reveals power law behaviour.

Figure 24b - Frequency of annual extinction rates from 1912-95 of the world's largest 100 companies in 1912. Uses the data from Figure 24a. (Ormerod 2005).

77
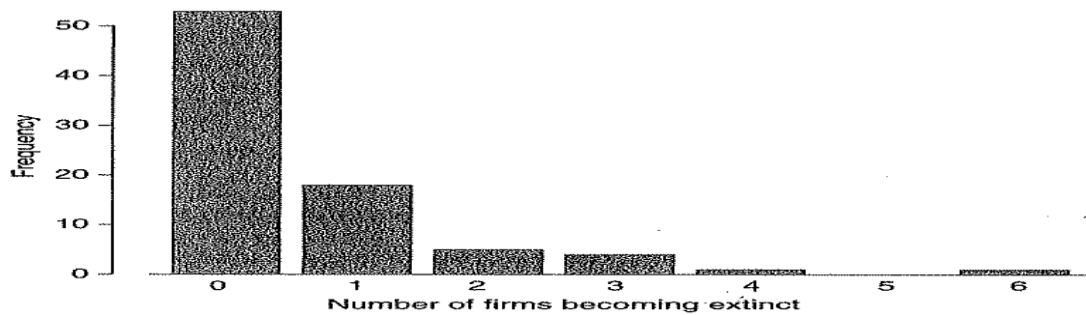
Figure 24c - Power law of the frequency and size of the extinctions from 1912-95 of the world's largest 100 companies in 1912. (Ormerod 2005).
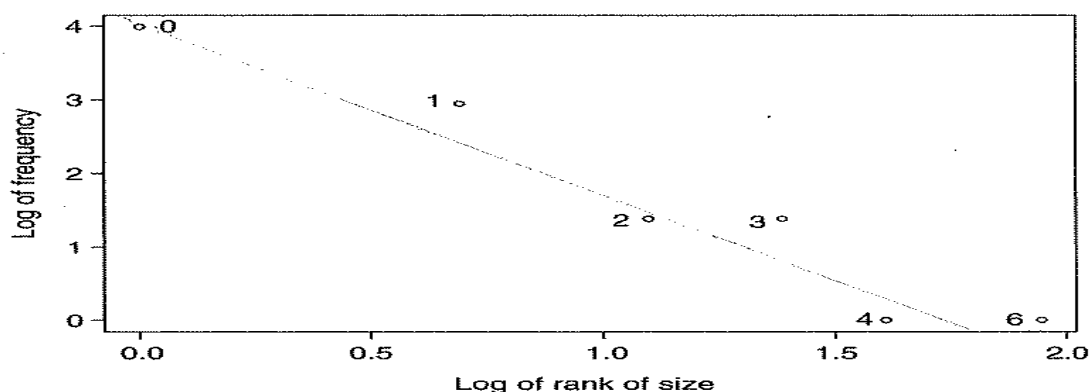
The precise mathematical relationship which describes the link between the frequency and size of the extinction of companies, for example, is virtually identical to that which describes the extinction of biological species in the fossil records. (Refer Figure 30a). Only the timescales differ. Ormerod questions how the patterns of failure, are so similar in biology and human organization when there is such a sharp contrast between the abilities to act with the conscious intent of improving one's prospects for survival.[1]

Traditional 'equilibrium' economic theory can not explain this non-linear effect

Ormerod believes that economic theory has a lot to say about what firms should do in order to succeed, but there is little, if anything, about why firms fail.[1]   More than 10 per cent of all economically active companies in the US become extinct each year (600,000 out of 6,000,000). It is a distinctive feature of companies, and any theory of the firm should attempt to explain it. Conventional economic theory can only do so by positing an endless supply of completely unexpected shocks, for otherwise the perfectly informed, rational decision-making firms could never die. They would live for ever. Mainstream economics likes 'exogenous' or external shocks, and uses phrases such as 'rational', 'perfect knowledge' and 'equilibrium'. Unfortunately the business cycle, the rate of growth of output (GDP) is constantly changing from year to year. Fluctuations have not disappeared.[2]

Both Schumpeter and Hayek were critical of conventional economics and both believed that it was precisely the complex interactions between the component parts that led to the pervasive existence of failure. And both realized that failure at the level of the individual component, the individual firm, could be beneficial to the system as a whole, since the firms that were dropping out were not as able to adapt to the changing conditions as the ones that survived.[3]

Hayek stressed the importance of dynamic, evolutionary change in the workings of human social and economic systems. Traditional theory however stressed, 'stability, order and

---

[3]   Ibid., Pp. 223- 234.

equilibrium' rather than 'constant change, evolution and disequilibrium', which is the experience of the real world.[1]

Complexity explanation for the power law effect

Power law behaviour at the level of the system as a whole arises from the ways in which the component parts of the system are connected. For example, the economy is characterized by its network of connections; of how, for example, firms are linked to each other, and whether the flow across any particular connection is positive or negative. Sometimes, it is the distribution of the connections themselves which follows a power law. Sometimes, it is the connections which give rise to power law behaviour of another feature of the system, such as the relationship between the size and frequency of extinction events in the fossil record.[2]

The developed economies of the west each have an extensive network of connections, especially the very large firms with each other. General Motors, for example, will almost certainly have dealings with software companies, computer manufacturers, banks and insurance companies, oil suppliers, producers of heavy industrial materials, and so on. Direct competitors are also connected to the company, since they have an impact on General Motors' fitness for survival and all its connected companies.[2]

Most of the time, small events, small shocks to the system, will only have small impacts, and large shocks will usually have big consequences. But the fact that we observe power-law behaviour in a system tells us that the system operates in a way that means that these relationships do not always hold. Sometimes, a very small event can have profound consequences, and occasionally a big shock can be contained and be of little import. It is not the power law itself which gives rise to these unexpected features of causality; rather it is the fact that we observe a power law in a system which tells us that causality within it behaves in this way. The conventional way of linear thinking, which postulates a link between the size of an event and its consequences, is broken.[2]

The relationship that describes the connection between the frequency and size of extinctions of biological species is the same as that which describes extinctions of companies. The timescales differ dramatically, but frequency and size are connected identically in both. This confronts us with a paradox. No biological species, with the exception of humanity, is able to anticipate the future and to plan its strategy accordingly. In reality, extinction is a pervasive feature for biological species, as it is for firms. Yet people in companies are able to think about strategy, they are able to make decisions which will effect the ability of the firm to survive, and still extinction is a pervasive feature.[2]

There are two interpretations of extinctions of firms. Traditional economics believes that extinctions are caused solely by shocks that are external to the system. Complexity says that extinctions of all sizes arise quite naturally from within the system itself, from the network of

connections between the individual agents and how they evolve over time. A variation to this, is where some complexity writers believe that shocks can be both internal and external rather than just external or just internal.[2]

Chapter 12 - Earthquakes

Gutenberg-Richter law of Earthquakes (Non-linearity)

The Gutenberg-Richter law for earthquakes is a power law and that means the distribution is scale invariant. Buchanan believes that there is nothing whatsoever about the large quakes to suggest they have an origin different from the small quakes. In the absence of any reason for suggesting that the really large quakes are special, the paradoxical implication is that what triggers small and large quakes is precisely the same. From this perspective, it does not make sense to look for special explanations for the largest earthquakes. They are no more special or unusual than the tiny shudders constantly rippling beneath our feet. It is absolutely crucial to realise that we cannot draw this conclusion from any mathematical form other than a power law. But a power law demands it. In view of the Gutenberg-Richter power law, it is extremely unlikely that the project of mega-quake prediction is possible. Indeed, the entire project of earthquake prediction may be fundamentally misguided, and practically impossible.[1]

'Equilibrium' and 'linear' approaches have failed to predict earthquakes

Per Bak is clear that scientists are poor at making earthquake predictions, and not for lack of effort. All kinds of approaches have been suggested, but none are consistent with the data. Any prediction theory of earthquakes must explain the Gutenberg-Richter law.[2]

The statistics for all quakes reveal a strikingly simple pattern. (Refer Figure 25). Between 1987 and 1996, for example, a large number of earthquakes hit southern California in the United States. In terms of energy, the data indicates that whenever you double the energy released by an earthquake, it is one quarter less frequent.[1] The Gutenberg-Richter law is surprisingly robust and does not vary significantly from region to region over time.[3] The Gutenberg-Richter law is a classic example of a simple law underlying what looks at first sight to be a complex system.[4]

[1] Buchanan, Mark (2000). *Ubiquity – The Science of history…Or Why the World is Simpler Than We Think.* Weidenfeld and Nicholson. London. England. Pp. 22-42.

[2] Bak, Per (1997). *How Nature Works – The Science of Self-Organized Criticality*. Oxford University Press. Oxford, U.K. Pp. 86-87.

[3]   Wikipedia - Nonlinear system. http://en.wikipedia.org/wiki/Nonlinear_system 19/12/2009

[4]   Gribbin, John (2004). *Deep Simplicity. Chaos, Complexity and the Emergence of Life.* Allen Lane. An imprint of Penguin Books. London, England. Pp. 138-143.

Figure 25
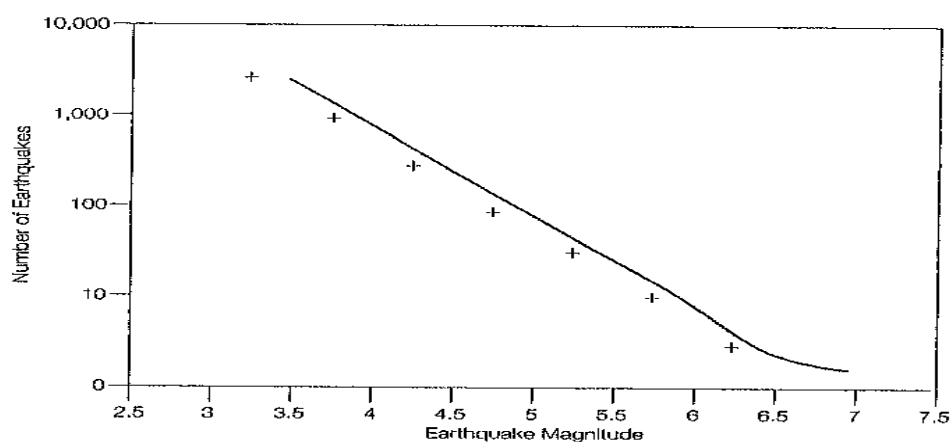
Power law curve for earthquakes in Southern California between 1987 and 1996 from Buchanan. (Beinhocker – 2007).

Logarithmic scaling

The Gutenberg/ Richter graph does not seem to offer much illumination regarding the question about what is typical when it comes to earthquakes. But in other ways it tells us a lot.[1] The Richter scale is logarithmic, so that an increase of one unit on the scale corresponds to an increase in the amount of energy released by a factor of 30; a magnitude 2 earthquake is 30 times as powerful as a magnitude 1 earthquake, a magnitude 3 earthquake is 30 times more powerful than a magnitude 2 earthquake (and therefore 900 times more powerful than a magnitude 1 earthquake), and so on.[4]

The number of earthquakes of each magnitude also obey a power law – for every 1,000 earthquakes of magnitude 5 there are roughly 100 earthquakes of magnitude 6, 10 earthquakes of magnitude 7, and so on. An earthquake of magnitude 8, a little smaller than the famous San Francisco earthquake of 1906, is 20 billion times more energetic than an earthquake of magnitude 1, which corresponds to the kind of tremor you feel indoors when a heavy lorry passes by in the street outside.[4]

Since earthquakes follow power laws, there is no typical size. Earthquakes occur across all size scales, but the bigger the quake, the rarer it is. Power laws, along with oscillations and punctuated equilibrium, are signature characteristics of complex adaptive systems.[5] Thus, it appears that one should not try to come up with specific explanations for large earthquakes, but rather with a general theory encompassing all earthquakes, large and small.[6]

---

[5] Beinhocker, Eric D. (2007). *The Origin of Wealth. Evolution, Complexity, and the Radical Remaking of Economics.* Random House Business Books. London, Great Britain. Pp. 178-179.

[6] Bak, Per (1997), op. cit., Pp. 14-18.

Scale invariant

A power law, with its scale-invariant form, implies that large events are just magnified copies of smaller ones, and arise from the same kinds of causes. Really big earthquakes are not triggered by special events, but are simply the natural if infrequent consequence of the overall critical organisation of the earth's crust, and its susceptibility to long-range chain reactions.[7]

Buchanan believes that when power laws apply, then words like 'normal', 'typical', 'abnormal' and 'exceptional' do not apply. Power laws always have this implication, whatever the subject being considered.[1] Large earthquakes occur more rarely than small earthquakes, but they are produced by essentially the same physical process as small earthquakes –and the power law tells you this even if you have no idea what that single physical cause of earthquakes is.[4]

This is quite different from the idea, still widely held, that large earthquakes require big triggers. If that were so, then it would be comforting, as you would wait a long time before the next big strike. The statistics and the power law in particular, say otherwise. Both large and small earthquakes occur at random, but importantly with different frequencies. Gribbin believes that an earthquake of any size could happen at any time in an earthquake zone, and just as the possibility of a tossed coin coming down heads is still 1 in 2 even if the last three tosses have all come down tails, so the chances of another earthquake like the event of 1906 were no less (and no more) in 1907 than they were in 1905. Power laws always mean that the thing being described by the law is scale invariant, so that earthquakes of any size are governed by exactly the same rules.[4]

Omori's law of aftershocks

Omori's law states that the rate of aftershocks decreases quickly with time. The rate of aftershocks is proportional to the inverse of the time since the main shock. Thus whatever the odds of an aftershock are on the first day, the second day will be ½ the odds of the first day and the tenth day will be approximately 1/10th the odds of the first day (when p is equal to 1). These patterns describe only the mass behaviour of aftershocks; the actual times, numbers and locations of the aftershocks are 'random', while tending to follow these patterns. Aftershocks follow the Richter law of scaling, so in summary there are more small aftershocks and fewer large aftershocks.[8]

Said another way, large earthquakes tend to cluster together in time, and the longer you wait without seeing one, the longer you will probably have to wait. The mathematical signature of this clustering is the distribution of 'waiting times' for the next large earthquake, after one large quake has taken place, a distribution that also follows a power law (Omori law) and is 2.6 – that means, waiting times of, say, two weeks turns out about 2.6 times less frequent

[7]   Buchanan, Mark (2000), op. cit., Pp. 66-88.

[8]   Wikipedia – *Aftershock*. http://en.wikipedia.org/wiki/Aftershock#Omori.27s_Law 07/05/2010.

than times of just one week, and the same goes for times of two months relative to one month, two years relative to one year, and so on.[7]

Clustering

Power law phenomenon such as earthquakes are inherently not periodic (i.e. don't follow regular cycles). The power law indicates that the longer you have waited since a large earthquake at a given location, the longer you can expect to wait. Earthquakes are clustered in time, not periodic. This applies to evolution, where the longer a species has been in existence (i.e. cockroaches), the longer we can expect it to survive. This is contrary to popular thinking. People assume that since an earthquake is due at a particular location, means that one will happen there soon.[9] Regularity should not be confused with periodicity.[6]

Careful statistical studies of earthquakes indicate that the longer some region goes without a quake, the less likely it is to suffer one in the near future. Earthquakes seem to strike in clusters. The best current method for predicting an earthquake according to Per Bak is probably to wait until one strikes, and then quickly predict another at that location.[1]

---

[9]   Bak, Per (1997), op. cit., Pp. 174.

Chapter 13 - Scientific progress and scientific citations

Scientific citations

Alfred Lotka reported in 1926 in his ground breaking discovery of the so-called Law of Scientific Productivity, which states that the distribution of the numbers of papers written by individual scientists follows a power law. That is, the number of scientists who have written k papers falls off for some constant α. In fact, this result extends to the arts and humanities as well. Both the in- and out-degree (numbers of times a paper has been cited and number of other papers a paper cites) distributions of the network follow power laws, a far reaching discovery.[1]

Lotka's law is now defined as the frequency of publication by authors in any given field. As the number of articles published increases, authors producing that many publications become less frequent. The formula is $1/n^a$ where 'n' is authors and 'a' is usually 2. For example, there are ¼ as many authors publishing two articles within a specified time period as there are single-published authors, 1/9 as many publishing three articles, 1/16 as many publishing four articles, etc. Though the law itself covers many disciplines, the actual ratios involved are discipline related. Lotka's law is believed to have application in other fields for example in the military for fighter pilot kills.[2]

In the 1950s Herbert Simon[1] showed that power laws arise when the 'rich get richer', when the amount you get goes up with the amount you already have. In sociology this is referred to as the Matthew effect, after the biblical edict, 'For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath.' (Matthew 25:29 – King James Bible).

Later studies showed that the network of citations between scientific papers followed a power law and that many citations receive further citations in proportion to the number they already have. This was called 'cumulative advantage' or as it is more commonly called 'preferential attachment'. Newman believes that preferential attachment is the most widely accepted explanation for the occurrence of power-law degree distributions in networks as diverse as the World Wide Web, social networks, and biological networks.[3]

Redner (1998) also independently discovered power law effects using two large databases of citations of physics papers. Redner investigated the citation frequency of 783,339 papers

---

[1] Newman, M.E.J. (2003). *The structure and function of complex networks.* arXiv:cond-mat/0303516v1 (cond-mat.stat-mech)   25 March 2003. Pp. 6-7 and 30.

[2] Wikipedia – *Lotka's Law*. http://en.wikipedia.org/wiki/Lotka%27s_law 07/05/2010

[3] Newman, Mark; Barabasi, Albert-Laszlo; Watts, Duncan J. (2006). *The Structure and Dynamics of Networks.* Princeton University Press. Princeton. New Jersey. U.S.A. Pp. 17-18.

published in 1981 and cited over 6 million times between 1981 and 1997. Redner found the power tail exponent roughly equal to -3. The data shows a 'scale-free network'.[3] Most papers were cited only a few times, but a few appeared in a large number of reference lists. In fact, about half the papers (368,110 papers) were never or hardly ever cited at all, a telling reminder of the modest nature of most research projects. Of course some research papers are highly influential.[4] (Refer Figure 26).

Figure 26
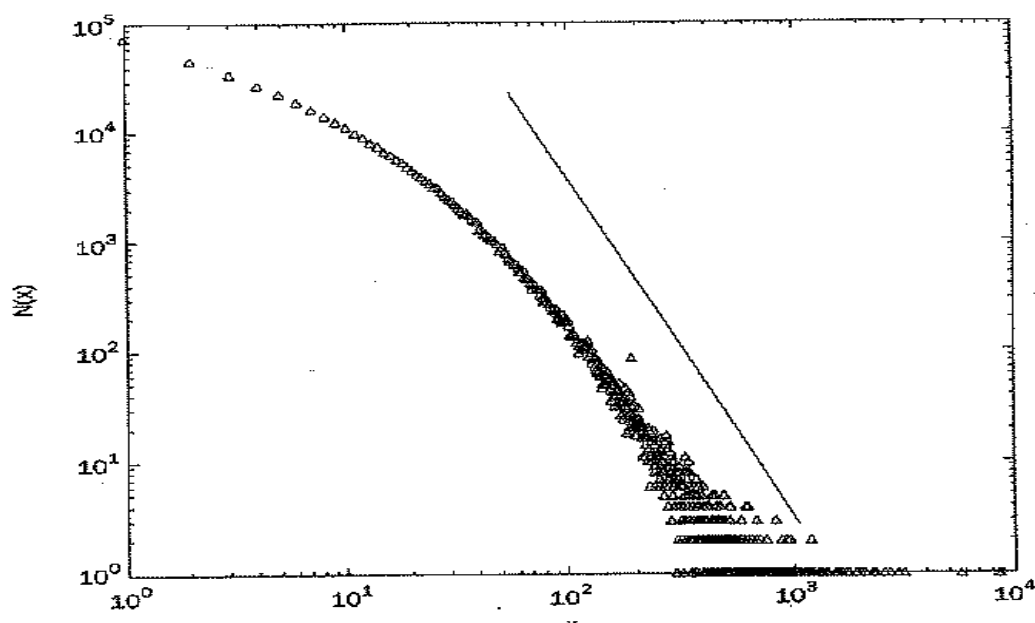
Power law curve by Redner of the distribution of research papers according to the number of citations they receive. (Buchanan – 2000).

For the papers that collected more than a hundred citations or so, Redner found that the distribution of citations follows a scale-invariant power law; precisely what we might expect if the network of ideas is organised, like the sand-pile, into a critical state. Redner discovered that the way highly cited papers become rare follows an extremely regular pattern: double the number of citations, and the number of papers receiving that many falls off by about eight. So there is no typical number of citations for a paper, and by extension, no typical magnitude for the reshaping in the network of ideas that any paper ultimately entails.[5]

Mathematical models by Newman of the scientific citation process predict a strong 'first-mover' effect under which the first papers in a field will, essentially regardless of content,

---

[4] Ball, Philip (2004). *Critical Mass – How one thing leads to another.* Arrow Books. London. United Kingdom. Pp. 302.

[5] Buchanan, Mark (2000). *Ubiquity – The Science of history…Or Why the World is Simpler Than We Think.* Weidenfeld and Nicholson. London. England. Pp. 180-184.

receive citations at a rate enormously higher than papers published later. This is confirmed by empirical evidence.[6]

The collaboration networks of scientists in biology and medicine, various sub disciplines of physics and computer science were analyzed by Newman, by using author attributions from papers or reprints appearing in those areas over a 5-year period from 1995 to 1999. They all form 'small world' networks, in which the average distance between scientists via a line of intermediate collaborators varies logarithmically with the size of the relevant community. Newman found that the three databases had a power law result with an exponential cut-off. The power law exponent was constant at about 2.5.[7]

Newman studied the distributions of both the number of collaborators of scientists and the number of papers they write, and found both follow power law distributions with an exponential cutoff. This cutoff may be caused by the finite time window used in the study.[7]

'Punctuated equilibrium' describes how science progresses

Thomas Kuhn described the development of science through periods of stable growth punctuated by revisionary revolutions. According to Kuhn the development of science is not uniform but has alternating 'normal' and 'revolutionary' phases. The revolutionary phases are not merely periods of accelerated progress, but differ qualitatively from normal science. Normal science does resemble the standard cumulative picture of scientific progress, on the surface at least.[8] Kuhn called the violent revolutionary periods of upheaval 'paradigm shifts'.[9] Discovery of new paradigms in science is the sudden 'punctuation', whilst 'normal' science is the gradual 'equilibrium'.[10]

Scientists like Albert Einstein, Isaac Newton and Charles Darwin are names we would match against the great, earth-shattering scientific revolutions. In a 1969 postscript to a new edition of The Structure of Scientific Revolutions, Kuhn emphasised that such revolutions need not have widespread consequences or involve fundamental ideas. A small sub-field of physics, say, or even a handful of scientists in a single research group, can experience a revolutionary change in the structure of ideas that form the basis of its work. If the ideas on which it has

---

6   Newman, M.E.J (2008). *The First-Mover advantage in Scientific Publications*. arXiv:0809.0522vl (physics.soc-ph) 2 Sep 2008.

7   Newman, Mark; Barabasi, Albert-Laszlo; Watts, Duncan J. (2006). *The Structure and Dynamics of Networks*. Princeton University Press. Princeton. New Jersey. U.S.A. P 221-5   M.E. J. Newman – The Structure of Scientific Collaboration Networks – Pp. 404-409, PNAS, January 16, 2001. Volume 98. No. 2.

8   Stanford Encyclopedia of Philosophy. *Thomas Kuhn.* Retrieved 07/08/2010. http:/plato.stanford.edu/entries/Thomas-kuhn/

9   Beinhocker, Eric D. (2007). *The Origin of Wealth. Evolution, Complexity, and The Radical Remaking of Economics.* Random House Business Books. London, Great Britain. Pp. Xi.

10   Kuhn, Thomas S. (1996). *The Structure of Scientific Revolutions*. Third Edition. The University of Chicago Press. Chicago. U.S.A.

depended increasingly fail to deliver results, even a small group may experience the same basic pattern of change. A major revolution does not have to be exclusively associated with an Einstein, Newton or Darwin. So in Kuhn's mind, what distinguishes a revolution from normal science is its tradition-shattering as opposed to tradition-preserving character. A revolution tears up part of the old network of ideas; normal science simply adds to it. Kuhn's discussion hints at the possibility that there may be no typical revolution; that the dynamics of science could be scale-invariant, and that the network of good ideas might, like the earth's crust, be poised in a critical state.[11]

Throughout all these phenomena, we make a natural but false distinction between the few big events and the background of the more numerous run-of-the-mill events. For example, we like to separate the five mass extinctions from the numerous smaller extinctions caused by evolution; or we like to separate the few great earthquakes from the numerous smaller quakes. In each of these examples, there are no exceptional causes. It looks as if something very similar may be true of science itself. Kuhn hinted that there could be small revolutions as well as large, and that both shared the same essential character as 'tradition-shattering' events. Redner's power law for citations is a kind of Gutenberg-Richter law for scientific upheavals, and implies that there is, in a deep sense, no true distinction between large and small scientific revolutions.[11]

Buchanan believes that scientific knowledge is like many other things poised in a critical state. If so, then scientists should expect the unexpected. For the fabric of ideas is organised so that the tiniest chance discovery may at any moment and without warning activate a domino-like chain of effects leading to a terrific revolution. Foreseeing such revolutions is next to impossible, for the ultimate consequences of any new idea depends not so much on its own inherent profundity as on where it happens to fall within the network of all scientific ideas. As seen through the lens of the critical state, great revolutions are not necessarily unique in terms of their causes. They are simply the expected large fluctuations of a system poised in a critical state.[11]

The scale-free power law for citations, means that the smallest revolutions are happening every day, may involve only a few scientists in specialised communities, and may be virtually invisible to almost everyone else. By contrast, the largest revolutions may wipe away much of science as we know it, and are liable to happen at any time, if the right idea pops up in the right place. So while chains of events in these systems may not be predictable, it is not the case that nothing is predictable. It is in the statistical pattern that emerges over many chains of events that we can hope to discover the laws for things historical.[11]

---

[11]   Buchanan, Mark (2000), op. cit., Pp. 176-184.

Chapter 14 - Species extinctions

Power law distributions are key to understanding evolution

Mark Newman believes that all mathematical extinction models need to take into account three key global properties of fossil records.[1] These are:

(1) The fossil record appears to have a power-law (i.e., scale- free) distribution of the size of extinction events, with an exponent close to 2.

(2) The distribution of the lifetimes of species also appears to follow a power law, with an exponent of about 1.7.

(3) The number of species per genus also appears to follow a power law with an exponent of about 1.5.

Extinction rates

Figure 27a shows a plot of the number of families of marine organisms becoming extinct in each geological stage since the start of life. The graph shows more variation than the five mass extinctions. To investigate the endogenous view of extinction, researchers have constructed plots of the distribution of the sizes of extinction events; non-uniformity in such a distribution might offer support for distinct mass and background extinction mechanisms.[2]

Figure 27a
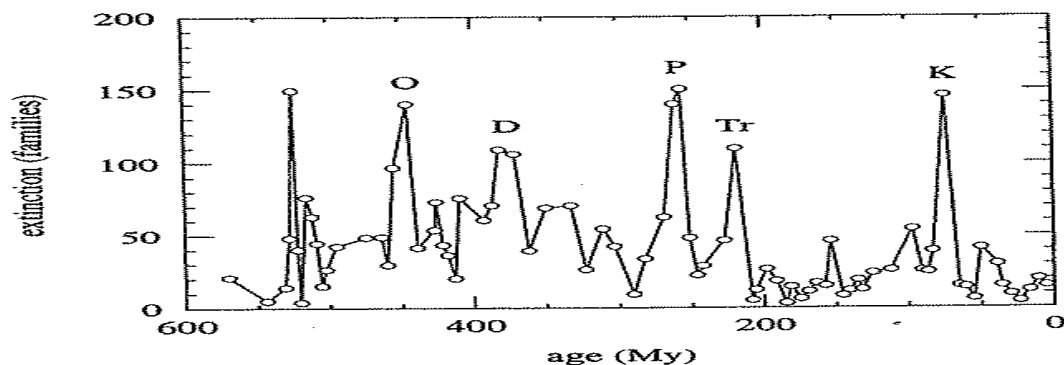
The number of families of known marine organisms becoming extinct per stratigraphic stage as a function of time. The positions of the 'big five' mass extinctions are marked with letters. (My = millions of years.) Data from Sepkoski. (Newman 2003).

Figure 27b is a histogram of the data in Figure 27a. Figure 27b is highly skewed and very far from a normal distribution. It shows no visible discontinuities, within the sampling errors, and

---

[1] Newman, M.E.J and Palmer, R.G. (2003). *Modelling Extinction.* A Volume in the Santa Fe Institute Studies in the Sciences of Complexity. Oxford University Press. New York. U.S.A. Pp. 81-83.

[2]   Ibid., Pp. 1-22.

therefore gives no evidence for any distinction between mass and background extinction events.

Figure 27b
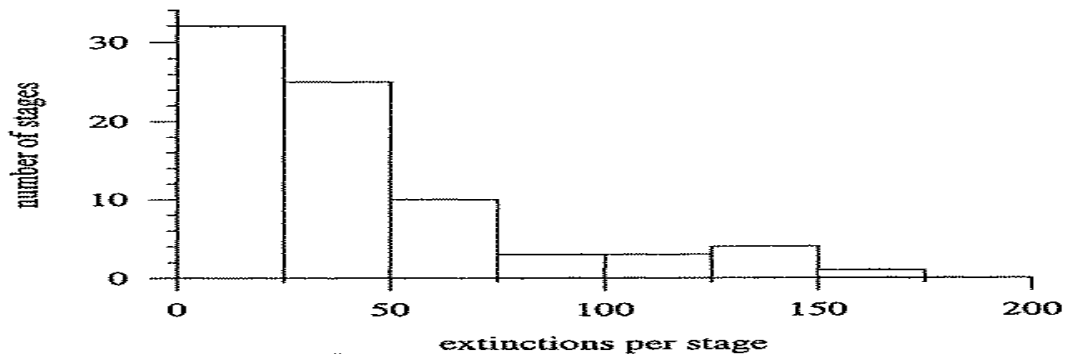
Histogram of the number of families of marine organisms becoming extinct. Data from Figure 27a - Sepkoski. (Newman 2003).

Figure 27c is the data from 27b replotted using logarithmic scales using the Zipf ranking method, which shows a straight line of a power law distribution, at least for large extinction events. Whilst the graph shows both a power law fit and an exponential one, Newman believes that the power law is a better fit with a power law exponent of approximately 2. [2] Newman's methodology is considered in greater detail in Chapter 17, 'Not all power law distributions are power law'.

Figure 27c
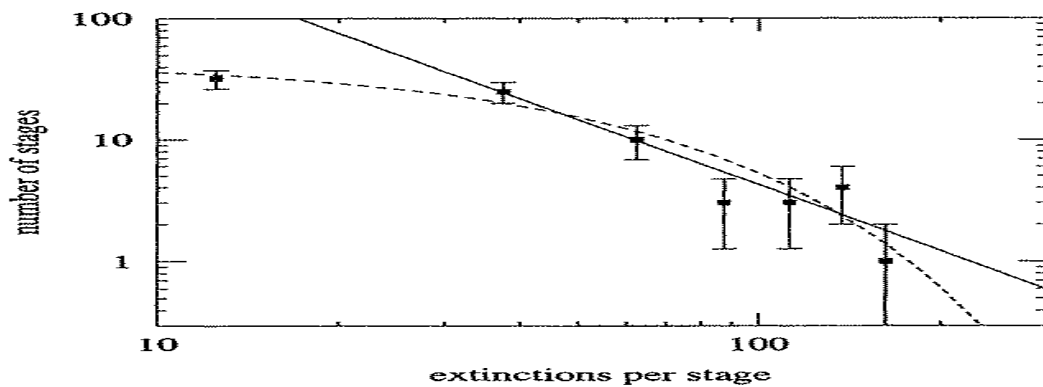
Power law of the data from Figure 27b replotted on logarithmic scale, with Poissonian error bars. The solid line is the best power-law fit to the data. The dashed line is the best exponential fit. (Newman 2003)

Species lifetimes

Another quantity which has been compared with the predictions of a variety of extinction models is the distribution of the lifetimes of groups of organisations (taxa). Figure 27d shows a log-log graph of the lifetimes of marine species in the Sepkoski database. The figure's axes are logarithmic and the solid and dotted lines represent respectively power law and

exponential fits to the data. At first glance it appears from this figure that the lifetime distribution is better fitted by the exponential form. However, the power-law fit in the figure is a fit only to the data between 10 and 100 million years. For this interval, the power law line matches the data quite well. Newman believes that both the very short and very long lifetimes are probably under-represented in the database because of systematic biases.[2]

Figure 27d
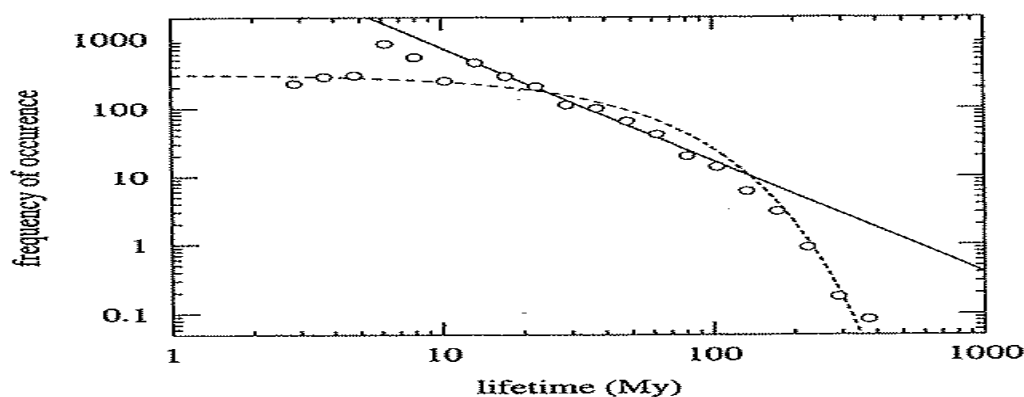
Power law of the frequency distribution of marine species lifetimes in the fossil record. The solid line is the best power-law fit to the data between 10 and 100 million years (My), while the dotted line is the best exponential fit to all the data. After Newman and Sibani. (Newman 2003).

First, since the appearance and disappearance of species are recorded only to the nearest stage, lifetimes of less than the length of the corresponding stage are registered as being zero and do not appear on the histogram. This means that lifetimes shorter than the average stage length of about 7 million years are underrepresented. The longest lifetime data is always slightly shorter (i.e. up to 7 million years) than it may have been, so this causes some underestimation in results. Overall, Newman believes the power law form is a reasonable conjecture. The exponent of the power law for the central region of the figure is measured to be approximately 1.7.[2]

The number of species per genus

As long ago as 1922, Willis noted that if one takes the taxonomy hierarchy of current organisms, counts the number of species in each genus, and makes a histogram of the number of genera for each value, then the resulting graph has a form which closely follows a power law. In Figure 27e Willis used species per genus of flowering plants. The measured exponent in this case is approximately 1.5. Others have subsequently extended these results to higher taxa, showing that the number of genera per family, families per order, and so forth, also follow power laws, suggesting that the taxonomic tree has a fractal structure, a result of some interest to those working on 'critical' models of extinction.[2]

Figure 27e
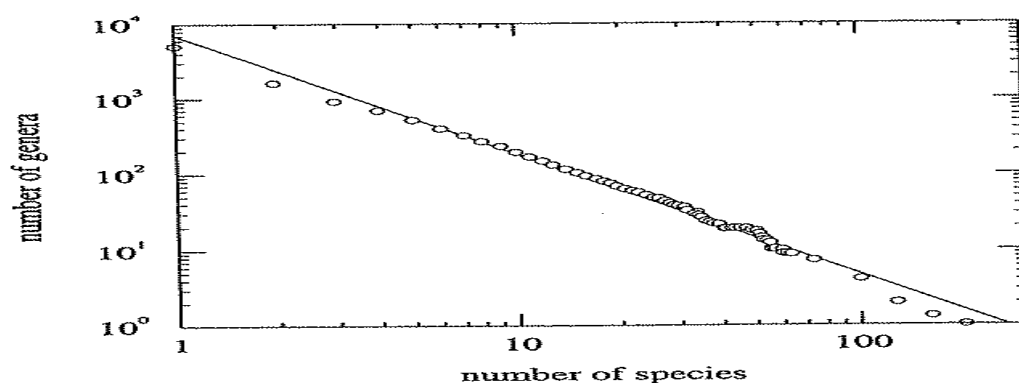
Power law of the number of species per genus of flowering plants. The solid line is the best power-law fit to the data. After Willis. (Newman 2003)

Extinctions triggered by both exogenous and endogenous reasons

Of the estimated one to four billion species that have existed on the Earth since life first appeared here, less than 50 million are still alive today. All the others became extinct, typically within about ten million years of their first appearance.[2]

There are two primary schools of thought about the causes of mass extinction. The traditional view is that extinction is the result of external stresses imposed on the ecosystem by the environment. There is good evidence for particular external/ exogenous causes for a number of major extinction events in the Earth's history, such as marine regression (sea level drop) and meteorite impact for the end of the Cretaceous. Other traditional explanations are also exogenous in nature, such as climate change, volcanism, tidal waves, etc. However, mass extinctions only account for 35 percent of the total extinction evidence in the fossil record at the family level, and for the remaining 65 percent we have no firm evidence favouring one cause over another.[2] The paleontological wisdom primarily leans in favour of the traditional abiotic approach.[1]

The alternative view is held by biologists and ecologists who support the idea that background extinction has biotic and internal/ endogenous causes – that extinction is a natural part of the dynamics of ecosystems and would take place regardless of any stresses arising from the environment. Those who hold this view are keen to use models which make predictions about trends based on the power law observations of the overall fossil patterns.[2]

For Per Bak, the 'external' force idea is essentially a 'linear' idea, since it assumes that a large effect must come from a large impact. In a non-linear world, there is no simple way to predict emergent behaviour. Self-organising criticality does not need a large change to make

a large impact, as it is not a linear effect. Bak is therefore strongly against searching for linear solutions (external forces) to non-linear systems.[3]

Stephen Gould suggested that the best way to look at it is to say that both endogenous triggers (often of small initial extent) and exogenous triggers (both great and small) can be in operation at the same time. The two approaches to causing triggers are complementary and do not exclude each other.[4]

Catastrophes – Punctuated equilibrium

The Darwin –Lyell viewpoint that extinction is a gradualist one is not supported by the evidence, which shows the regular occurrence of mass extinctions in the fossil records.[2]

Punctuated equilibrium was proposed in 1972 by Niles Eldredge and Stephen Gould, based on a study of fossil records suggesting that the evolution of individual species takes place in well-defined steps (the punctuation marks) separated by long periods of stability (equilibra or stasis). Extinctions occurred in bursts.[5]

Some regard the mass extinctions simply as a set of unrelated events. Others have sought a single mechanism to explain them all: the bombardment of the earth by comets, the destruction of shallow seas when continents collide and merge, a shortage of oxygen, a dramatic climate change, or a huge meteorite (favoured for killing the dinosaurs sixty-five million years ago).[5]

Mass extinctions occurred 65 million years ago and in a geological blink of the eye, the dinosaurs and 75 per cent of all other species suddenly went extinct. Other mass extinctions occurred 440, 365, 250, and 210 million years ago. These form the five big extinctions. Of course not every extinction is part of a mass extinction.[6]

Just because a meteorite (i.e. a 10 kilometre diameter meteorite that hit the Yucatan Peninsula of Mexico[7]) is the favoured explanation for the mass extinction of 65 million years ago, doesn't mean that meteorites are the main reason for mass extinctions. For example, if meteorites are an important trigger, then where are the meteorite craters that fit the other four of the five mass extinctions, being 210, 250, 365 and 440 million years ago?[6]

---

[3] Bak, Per (1997). *How Nature Works – The Science of Self-Organized Criticality*. Oxford University Press. Oxford, U.K. pp. 130-131.

[4] Gould, Stephen Jay (2007). *Punctuated Equilibrium.* The Belknap Press of Harvard University Press. Cambridge, Massachusetts. U.S.A. Pp. 235.

[5] Coveney, Peter and Highfield, Roger (1995). *Frontiers of Complexity. The search for Order in a Chaotic World.* Fawcett Columbine. New York. U.S.A. Pp. 232-234.

[6] Buchanan, Mark (2000). *Ubiquity – The Science of history…Or Why the World is Simpler Than We Think.* Weidenfeld and Nicholson. London. England. Pp. 102-112.

It seems that not all meteorites are equal in impact. For example, researchers in 1998 uncovered evidence of a huge crater 100 kilometres across in northern Siberia which was created at precisely the same time as an 85-kilometre hole was created at the mouth of Chesapeake Bay in the USA. Both were made 35 million years ago when comets smashed into the earth during a comet shower in the solar system. The fossil record shows absolutely nothing unusual at the time.[6]

Power laws come from systems that are in a self-organised critical state

Gribbin believes that what the fossil record seems to be telling us is that extinctions happen on all scales, all the time, and that (like earthquakes) an extinction of any size can happen at any time. Some extinctions might be triggered by meteorite impacts; some might be triggered by Ice Ages. But the other lesson from what we know about power laws and 1/f noise is that you do not need a large trigger in order to set off a large event. An extinction of any size might be set off by a trigger of any size. What matters is that we are dealing with a complex system – life on Earth – which is self-organising, feeding off a flow of energy, and exists at the edge of chaos.[7]

Per Bak suggests self-organizing criticality. Bak argues that life is a dynamic system, that far from ever existing in a stable steady state (equilibrium), organizes spontaneously into a characteristic and much more precarious critical state. Life does not evolve gradually but intermittently, with long periods of inactivity or stasis, interrupted by spurts of change characterised by mass extinctions and the emergence of new species. What is significant about this idea is that it does not require extinctions to originate through some 'external cause'. If life organises into a critical state, catastrophes, no matter how large, can be an intrinsic, self-organizing feature of evolution requiring no external causes.[5]

Using highly idealised computer simulations of evolution, Bak and colleagues found that during periods of frenzied evolutionary activity, the average fitness of species is low, as they relentlessly mutate in search of better fitness. The fitness of these various species is also low during mass extinctions, and high during periods of stasis associated with low evoluntary activity. Thus, in their work, 'survival of the fittest' did not mean evolution to a state where every species is well off: on the contrary, individual species are barely able to hang on – like the grains of sand in the critical sand pile or surviving a 'Red Queen effect'.[5]

According to their model, at the critical point all species influence each other. In this state they act collectively as a single 'metaorganism', many sharing the same fate. This is highlighted by the very existence of large scale extinctions. One corollary is that species with a high degree of complexity and many biological connections and dependencies, resulting from such factors as food chains, predator-prey and parasite-host relationships, are more sensitive to

[7]   Gribbin, John (2004). *Deep Simplicity. Chaos, Complexity and the Emergence of Life.* Allen Lane. An imprint of Penguin Books. London, England. Pp. 151-7.

fluctuations that perturb the dynamics and thus are more likely to be part of the next avalanche to extinction.[5]

The same power-law results occur with different sampling approaches

Jack Sepkoski's data (Figure 28) when plotted as a histogram (Figure 29) in four million years long intervals, found that just over 90 per cent of the extinctions occur in just under 50 per cent of the intervals. This is not a uniform distribution. Plotting each four million year long interval has fluctuated over the past 600 million years. The overall fluctuations of the graph look fairly random, with long intervals of relatively little variation, just a few species going extinct from time to time, punctuated by dramatic events in which large numbers of extinctions occur – the mass extinctions.[7]

Figure 28
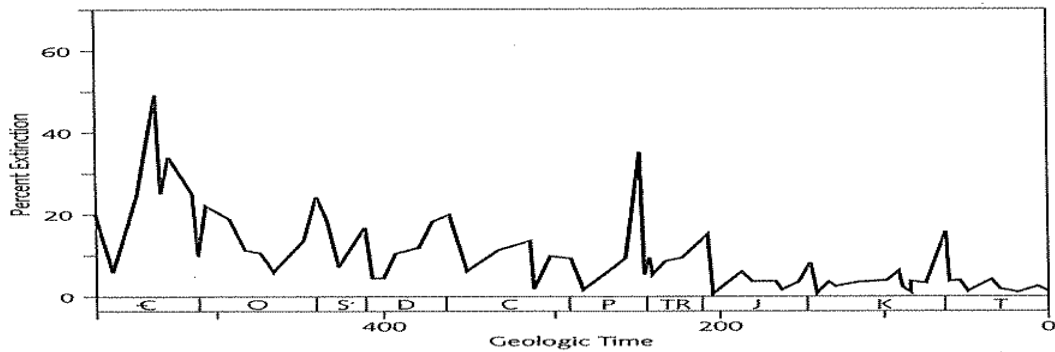
Biological extinctions over the last 600 million years, from Sepkoski. The curve shows the estimated percentage of families that went extinct within intervals of approximately 4 million years. (Bak – 1997).

Figure 29
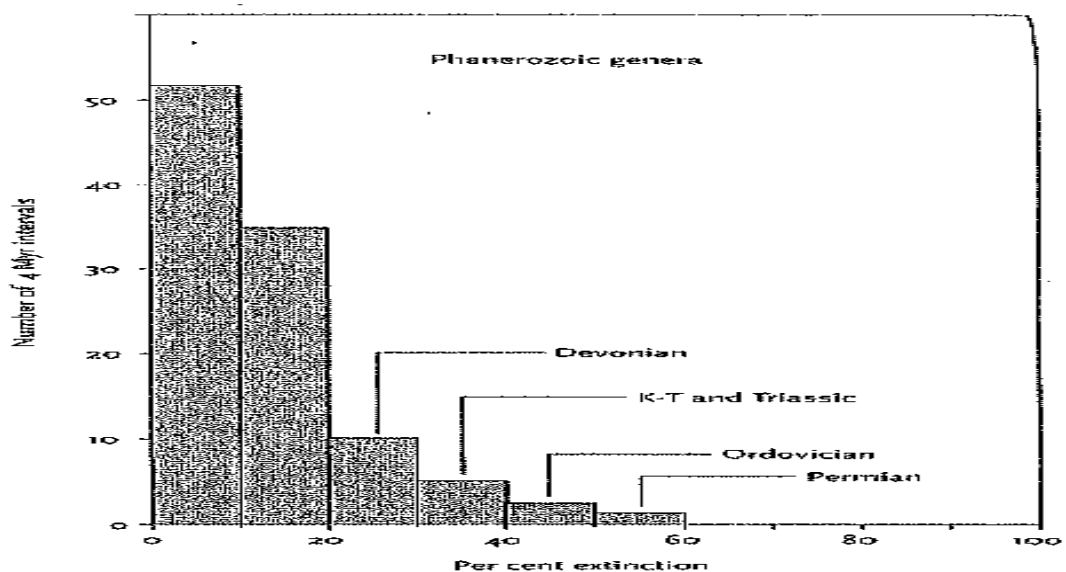
Biological extinctions – Histogram showing the number of intervals, each of four million years long, during which the indicated amount of extinctions occurred. Just over 90 per cent of the extinctions occur in just under 50 percent of the intervals, from Sepkoski. (Gribbin – 2004).

95

Like Sepkoski, Raup showed that the frequency with which we see extinctions of any given scale falls away with the square of the size (power law), for example, an extinction period in which twice the percentage of species per million years became extinct as in another event is four times less likely to be observed (four being two squared). A period in which three times as many become extinct is nine times less likely to take place, and so on.[8]

Figure 30b gives an illustration of the power law connection between frequency and size using the same Sepkoski data in Figure 28, which is plotted over time of the total of seventy-seven observations, each corresponding to a particular geological era. Forty-two of the observations experience an extinction rate of less than 3 per cent per million years, so the frequency is forty-two. We see an extinction rate of between 0 and 3 per cent forty-two times in the fossil record. Figure 30a is the histogram form of this data, which fits in sequence between Figure 28 and Figure 30b. [8]

The power law is the inverse square of the size. The label '0-3' refers to the percentage of total species becoming extinct every million years in the geological eras of the data set, and refers specifically to percentages between 0 and 3. Similarly, '3-6' refers to values between 3 and 6 per cent, and so.[8]

Figure 30a
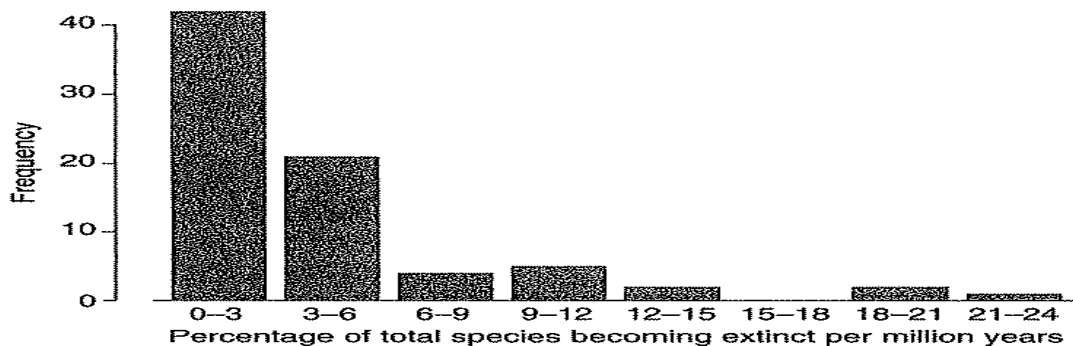
Frequency of extinction rates of biological species in geological eras since 550 million years before present. (Ormerod 2005).

---

[8] Ormerod, Paul (2005). *Why Most Things Fail – Evolution, Extinction and Economics*. Pantheon Books, New York, U.S.A. Pp. 160-200.

Figure 30b
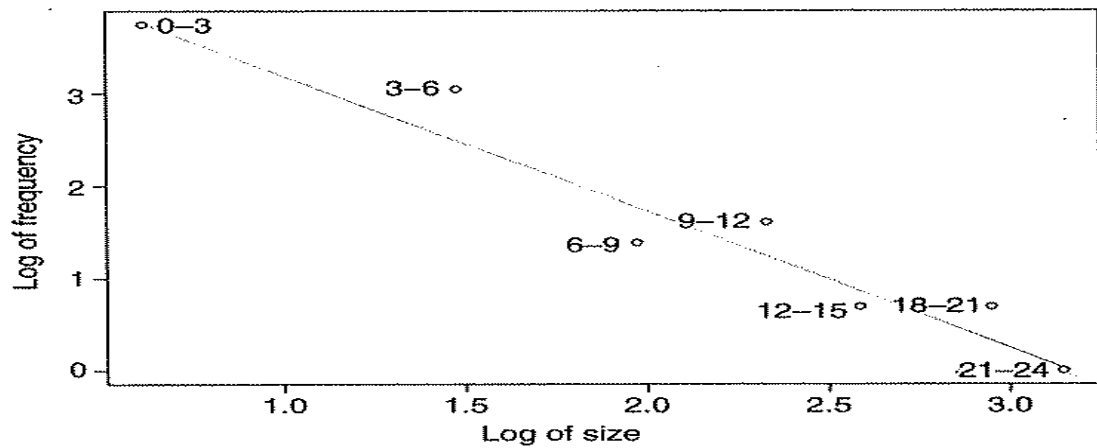
Power law of the frequency of extinction events in the fossil records against the (natural) log of the size of the extinction. Size is defined as the percentage of all species becoming extinct per million years, and this data is grouped into bands of three, i.e., all periods when the extinction rate was between 0-3 are in the '0-3' observations, and so on. (Ormerod 2005).

The striking feature of Figure 30b is that it follows a power-law distribution – the frequency of extinction falls much faster than the size of the extinction. Ormerod believes that this discovery of the empirical relationship between the size and frequency of extinctions is a major advance in how we need to think about evolution. He notes that it is strikingly similar to the annual extinction rate of the world's top 100 companies over the last 100 years.[8] (Refer to Figure 24c). These findings provide an important starting point for modellers to apply as the control data when modelling for the future or for different policy scenarios.

Chapter 15 - War

Richardson power law of war

In the 1920s, British physicist Lewis Richardson studied eighty-two wars that had flared up between 1820 and 1929. (Refer Figure 31). Richardson used deaths to judge the size of war.[1] Richardson analysed conflicts from individual murders (magnitude 0 conflict) to the two world wars (magnitude 7).[2] He found a power law; every time you double the number of deaths, the wars of that size become four times less common. This is precisely the same as the Gutenberg-Richter law, and we can draw similar conclusions. There is no typical size for a war, and no sensible division of wars into small skirmishes and major conflagrations. All wars whether over territory, religion or nationalism fall on a smooth curve that implies that the initial causes of all wars are probably the same. Later war historians like Jack Levy have used the number of casualties divided by the world population at the time. The power law still holds. The precise pattern becomes 2.62 times less frequent every time you double the number of deaths.[1]

Figure 31
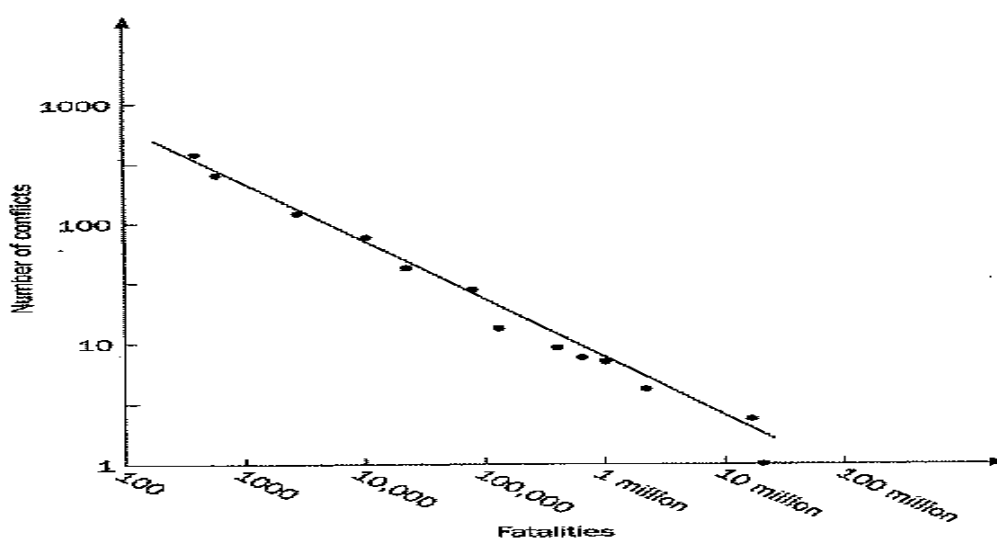
Power law curve for War - The distribution of deadly conflicts according to the number of people killed according to Richardson. (Buchanan – 2000).

Cederman believes Richardson's finding that the severity of interstate wars is power-law distributed belongs to the most striking, most accurate and most robust empirical regularities in world politics. Yet, in his view this is a regularity in search of a theory, as conventional theories on war can not account for the power law effect.[2]

---

[1] Buchanan, Mark (2000). *Ubiquity – The Science of history…Or Why the World is Simpler Than We Think.* Weidenfeld and Nicholson. London. England. Pp. 190-193.

[2] Cederman, Lars-Erik (2002). *Modelling the Size of Wars – From Billiard Balls to Sandpiles.* August 19, 2002. Department of Government, Harvard University. Mass. 02138, U.S.A. Pp. 1-20.

Predicting war is impossible

Levy notes that in spite of the importance of war and the enormous human and economic resources devoted to its many aspects, our understanding of war remains at an elementary level. There is no widely accepted theory of the causes of war and little agreement has emerged on the methodology through which these causes might be discovered. Instead, the literature is characterised by a proliferation of competing and often contradictory theories.[3]

Levy found that once a war (or a series of wars) is over, neither its incidence nor its seriousness has any impact on the likelihood of war in the period immediately following, regardless of its characteristics or frequency of occurrence and regardless of the total amount of war in a given period. These findings are valid for Great Power wars as well as for interstate wars involving the Powers over the last five centuries of the modern Great Power system.[4]

Particularly interesting are the implications of Levy's findings for the popular war-weariness hypothesis. Not only are most of the relevant test statistics relatively weak and statistically insignificant, but they tend to run in a direction contrary to that implied by the hypothesis. Levy found that neither war nor a series of wars retards subsequent war (or Great Power war). The severity of war has the smallest (rather than the largest) impact, and periods of the most destructive war (as measured by severity) tend to be followed by more (not fewer) Great Power wars.[4] Equally, there have been several empirical studies of the war contagion hypothesis. These consistently found that the outbreak of war is random rather than contagious, regardless of whether we look at short term contagion or long-term periodicity.[5]

Levy found that the only statistically valid trend that came out of all his analysis was that the frequency of war is inversely related to its seriousness. Wars have generally been either frequent but limited or infrequent but more serious; only rarely have they been both frequent and serious or infrequent and limited.[6]

Civilisation operates in a self-organised critical state

Ball suggests that civilisation operates in a self-organised critical state. His idea is that conflicts and wars are the result of tensions that hold international relations poised at the brink of catastrophe. As a consequence, wars of all magnitudes, from minor skirmishes to global

---

[3]   Levy, Jack S. (1983). *War in the Modern Great Power System 1495-1975*. The University Press of Kentucky. Lexington, Kentucky, U.S.A. Pp. 1.

[4]   Ibid., Pp. 166-167.

[5]   Levy, Jack S. (1982). *The Contagion of Great Power War Behaviour, 1495-1975*. American Journal of Political Science, Vol.26, No. 3. (Aug., 1982), pp. 562-584. University of Texas Press. U.S.A.

[6]   Levy, Jack S. and Morgan, T. Clifton (1984). *The Frequency and Seriousness of War: An Inverse Relationship?* The Journal of Conflict Resolution, Vol. 28, No. 4. (Dec., 1984), pp. 731-749. Sage Publications, Inc. U.S.A.

conflagrations, are inevitable.[7] Coveney believes that conflicts and social unrest tend to spread like avalanches.[8]

Like every other power law, the evidence is 'emergent' from the empirical evidence of the Richardson power law graph, i.e. the graph displays the relationship between the number of conflicts and their size, measured according to the number of fatalities that resulted. Wars which left only a few hundred dead and world wars which have killed millions, all follow a single power law.[7]

The power law implies that a war, when it starts out, does not 'know how big it will become' – nor does anyone else. The power-law, suggests that the world's political and social fabric tends to be organised on the very edge of instability, and in such a way that wars spread in a special way, so that their ultimate extent is nearly impossible to foresee. The scale free character of the power law suggests that at the outset of a war there are no obvious clues about how large it will become.[1]

Once a complex system is at the critical point and there is tension, small events can have disproportionate consequences. For example, the First World War started from the assassination of the Austro-Hungarian Archduke Franz Ferdinand.[7] It was hardly the first time a member of a ruling family was killed by an assassination and when they do occur they do not usually signal a reason for war. This is an example of the 'Butterfly' effect. In this case the flapping of the butterfly did cause a great storm, at other times the butterfly's flapping just caused it's wing to flap – no more and no less.

Modern severity/ frequency data

Cederman uses updated data from the Correlates of War Project (1998). In this project the focus is restricted to interstate rather than intrastate wars. Using double logarithmic scales, Figure 32 plots the severity of interstate war between 1820 and 1997. It is a power law since the fit is linear. While the data points in the lower right-hand corner correspond to the world wars, the vast majority of all other wars reach considerably lower levels of severity, though without straying very far from the estimate line. The slope estimate of -0.41 implies that a ten-fold increase of war severity decreases the probability of war by a factor of 2.6. These findings compare with those of Jack Levy's compilation of European great power wars from 1945 to 1965 which also had a straight line in a log-log diagram but with a steeper slope of -0.57.[2]

---

[7] Ball, Philip (2004). *Critical Mass – How one thing leads to another.* Arrow Books. London. United Kingdom. Pp. 303-304.

[8] Coveney, Peter and Highfield, Roger (1995). *Frontiers of Complexity. The search for Order in a Chaotic World.* Fawcett Columbine. New York. U.S.A. Pp. 183-188.

Figure 32
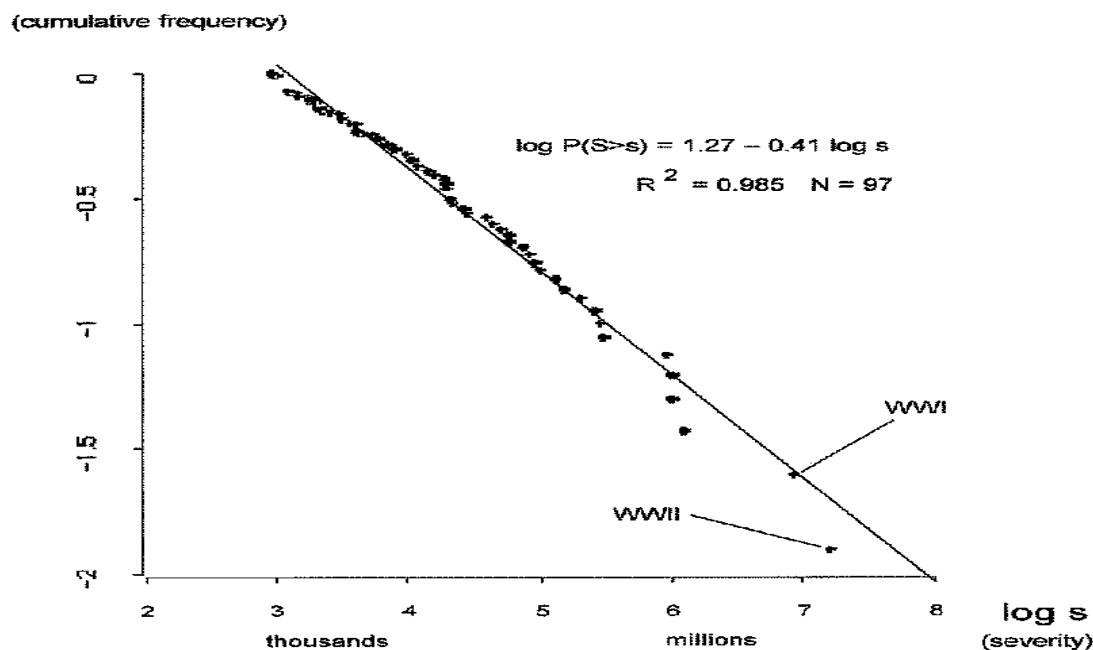
Power law curve for War - The cumulative frequency of severity of interstate wars, 1820-1997 according to COW data. (Cederman – 2002).

Given these strong results it may seem surprising that so few scholars have attempted to account for what seems to be an important empirical law. In fact, the situation is not very different from the economists' failure to explain the power law governing the distribution of city sizes, also known as Zipf's law. Using urban population data from many countries, researchers have established that the rank of city size typically correlates astonishingly well with city size. Cederman cites the economist Paul Krugman (1995) who noted that at this point we have to say that the rank-size rule is a major embarrassment for economic theory as it is one of the strongest statistical relationships we know, but lacks any clear basis in theory.[2]

Whilst International relations theory can not explain this power law effect, natural scientists have been studying self-organised criticality (SOC), which follows the ideas of Per Bak and others. What is unique about SOC systems is that they do not have to be carefully tuned to stay in the critical point where they generate the scale-free output responsible for the power laws. They do it automatically, i.e. they are 'self' organised. The non-equilibrium master metaphor is the sandpile. Whereas macro-level distributions emerge as stable features of scale-free systems, at the micro-level, such systems exhibit a strong degree of path dependence. To use the sandpile as an illustration, it matters exactly where and when the grains land. This means that point prediction often turns out to be futile, as exemplified by earthquakes. This does not mean, however, that no regularities exist. In particular, it is important to distinguish complex self-organized systems of the SOC kind from mere chaos, which also generates unpredictable behaviour.[2]

The stress on SOC encourages us to search for scale-invariant explanations. This means rather than treating large wars, such as world wars, as qualitatively distinct events that require separate explanations, that to get to a unified theory you need to explain all wars regardless of their size.[2]

Cederman found that technological change and contextual activated decision-making go a long way towards explaining why power laws emerge in geopolitical systems. Without these mechanisms, it becomes very hard to generate scale-free war-size distributions. Cederman's computational modelling of this regularity strengthens his confidence in the conjecture that interstate warfare actually follows the principles of self-organised criticality. If the SOC conjecture holds, then the logic of SOC casts doubt on international relation's static equilibrium theories as blue-prints for systematic theorising. If wars emanate from disequilibrium processes, then the narrow focus on equilibrium based theories is misguided.[2]

After all, scholars do make seismic analogies when reviewing history. For example, it is often stated that the end of the Cold War was caused by a series of geopolitical earthquakes.[2] This could equally be called the 'punctuated' part of the 'punctuated equilibrium' effect of history.

Cederman warns that care needs to be taken not to make simplistic analogies. For example, in the seismic analogy, he hopes that war is avoidable due to democratic security communities that can emerge, unlike the actions of the continental plates. Whereas some areas of the world are prone to frequent outbreaks of interstate violence, in others, catastrophic events are virtually unthinkable. Cederman suggests that the only safe way of managing security affairs is to transform the balance of power into a situation of trust, which is what happened between France and Germany in the last half century. While nuclear calamities would further vindicate Richardson's law, there would be few people around appreciating the advances of social science should the ostensibly 'impossible' turn out to be just another huge low-probability event.[2] Whilst I hope Cederman is right, I note that there is strong empirical support for the power law effect for war. This can only bode badly for the future, afterall there are always potential rogue states like Iran or Pakistan which are nuclear armed and ambitious. Arguably, the history of humanity has been as much the history of war and struggle as it has been the search for peace.

Insurgency attacks follow a power law pattern of timing and casualties

The power law roughly predicts the number of people that will be killed during a certain time period. Johnson and his team found that it's easier for enemy combatants to kill 10 people than 100. In fact, it's 316 (316 is 10 to the power of 2.5) times easier; therefore, smaller attacks are more likely than larger ones because they are less difficult to execute. Like earthquakes and stock volatility, violent attacks typically occur in 'bursts', such as multiple bombings in one city on a single day, and even at a certain time. Johnson believes you can't

predict what will happen tomorrow, but you can say what, statistically, will happen over a month. To many strategists and commanders, the timing of bursts of attack could suggest a coordinated effort. To Johnson and his colleagues, however, this suggests that enemy combatants saw a good opportunity and seized on it at the same time with little or no communication among rival groups. For insurgency groups do not operate in the traditional command and control style.[9]

Applying this to actual ongoing wars

The wars in Iraq and Afghanistan have been compared with the longer term guerrilla war in Colombia which is taking place against a back-drop of drug trafficking and Mafia activity. The origins, motivations, locations and durations of these conflicts are very different. However, the Johnson study found identical patterns of violence. (Refer Figure 33). Not only have the wars in Iraq and Colombia evolved to yield a power-law behaviour, but this behaviour is currently of the same quantitative form as the war in Afghanistan and global terrorism in non-G7 countries. The Johnson study findings suggest that the dynamic evolution of these various examples of modern conflict has less to do with geography, ideology, ethnicity or religion and more to do with the day-to-day mechanics of human insurgency. The respective insurgency forces are effectively becoming identical in terms of how they operate.[10]

Figure 33
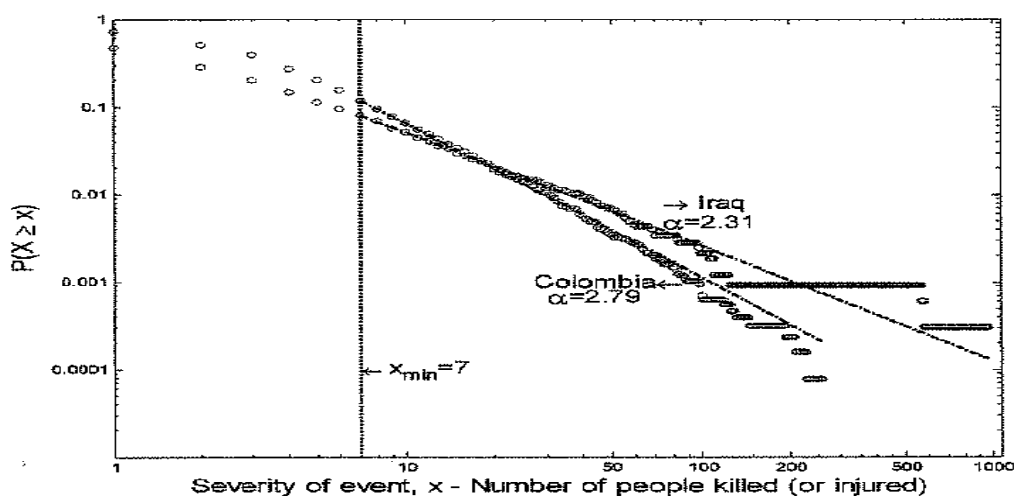
Power law for War - The distribution of deadly conflicts according to the number of people killed in Iraq (5/01/03 to 10/23/05) and Colombia (1988 to 2004). (Johnson et al. – 2005).

---

[9] Gilbert, Natasha (2009). *Modellers claim wars are predictable – Insurgent attacks follow a universal pattern of timing and casualties*. December 16, 2009.

www.scientificamerica.com/article.cfm?id=modellers-claim-wars-are-predi&print=true

[10] Johnson, Neil F.; Spagat, Mike; Restrepo, Jorge A.; Becerra, Oscar; Bohorquez, Juan Camilo; Suarez, Nicolas; Restrepo, Elvira Maria; and Zarama, Roberto. (2005) *Universal patterns underlying ongoing wars and terrorism.* http://xxx.lanl.gov/abs/physics/0506213 Oxford, London and Bogota, Colombia. Pp. 2-8.

Casualty numbers in global terrorist events, from 1968 to the present, obey power laws of α = 2.5 for non-G7 countries and 1.71 for G7 countries. For the wars in Iraq and Afghanistan, as well as the conflict in Colombia, each has a power law coefficients close to 2.5.[10] (Refer Figure 34).

Figure 34
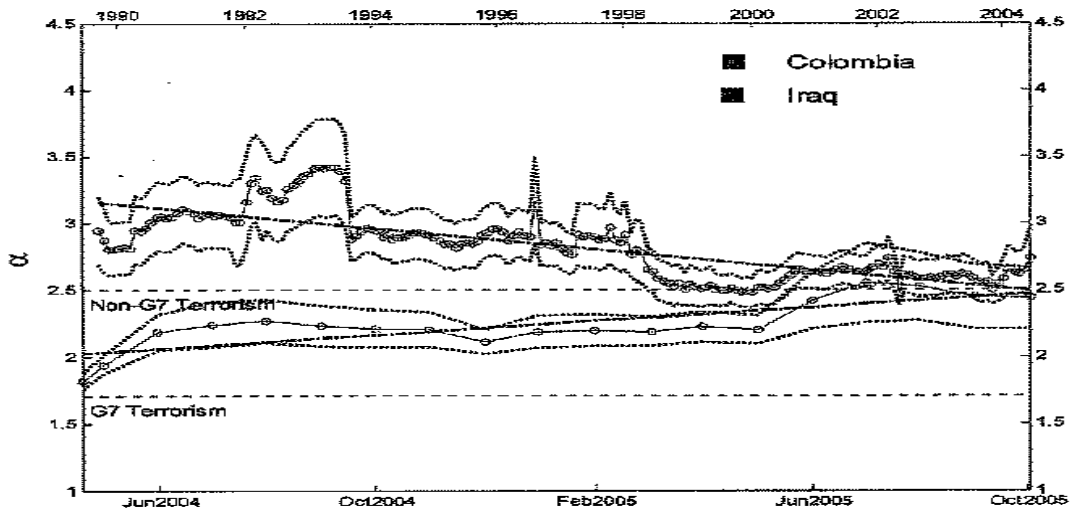
Power law for War - The variation through time of the power law coefficient for Iraq and Colombia. The graph includes a 95% confidence band around the two sets of data. (Johnson et al. – 2005).

As well as observing power-law behaviour for the time-aggregated data, it can also be seen over smaller time windows, i.e. α values as a function of time. At face value other wars such as Casamance (Senegal) (1982-2005), Indonesia (1996-2001), Israel (2000-2002) and Northern Ireland (1969-2001) have an average α between 2.4 and 2.5. The data covered nearly 55,000 separate acts of violence.[10]

Chapter 16 - Various networks

There are numerous types of network[1]. These include the following:

Type of network	Examples
Social networks	Film actors, company directors, science co-authorships, telephone call graph, email messages, email address books, human sexual contacts.
Information networks	WWW nd.edu, citation networks, word co-occurrence.
Technological networks	Internet, power grid, train routes, software packages, electronic circuits.
Biological networks	Metabolic networks, protein interaction, marine food web, neural network.

Many of the most important world networks – economic, political systems, ecosystems, and so on – are poised perpetually on the very edge of instability and tumultuous upheaval. As a consequence, it is something akin to a universal law of nature that the course of history must necessarily be punctuated – and quite frequently – by seemingly inexplicable upheavals. The character of history will therefore be long periods of relative calm and gradual change to be punctuated by staggeringly overwhelming events that totally reshape the social and political landscape.[2]

This chapter looks at a few of these power law affected systems. For example, movie-actor networks, social networks, the Internet, the World Wide Web, traffic jams, river networks and sexual partners.

Movie-actor network follows a power law

Gene Stanley and his colleagues looked closely at the power law in the movie-actor network, which seems inexorably to reward a few privileged individuals with ever better connections, and they found that this effect begins to falter at the high extreme. (Refer Figure 35). The best connected actors have rather fewer links than the power law would predict. The upper limit is set because actors have limited time and finite careers. There are also limits on most things, for example, on old scientific papers, volumes airports can handle etc.[3] Networks of movie actors therefore follow a power law distribution (i.e. a straight line in a log-log graph) with an exponential cut-off at the top end.[1]

[1] Newman, M.E.J. (2003). *The structure and function of complex networks.*
arXiv:cond-mat/0303516v1 (cond-mat.stat-mech)   25 March 2003. Pp. 7-13.

[2] Buchanan, Mark (2002). *Nexus – Small Worlds and the Groundbreaking Science of Networks.*
W.W. Norton and Company. New York. U.S.A. Pp. 20 and 83-85.

[3] Ball, Philip (2004). *Critical Mass – How one thing leads to another.* Arrow Books. London. United Kingdom. Pp. 479-501.

Figure 35
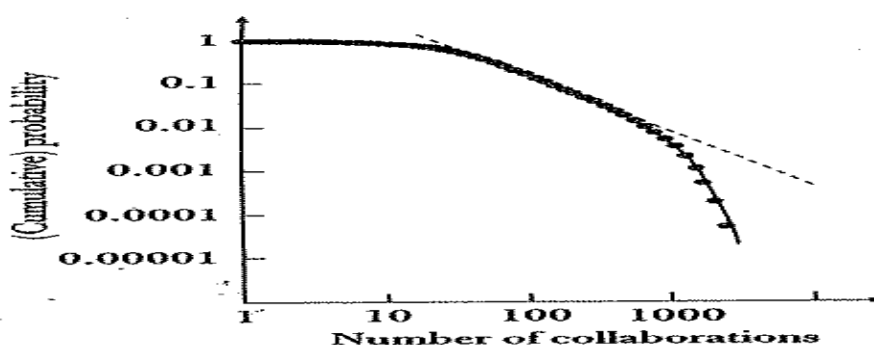
Power law curve for connections between actors – The most highly connected movie actors do not quite follow the power law that characterizes the rest of the network as there appears to be an upper limit to the number of collaborations an actor can attain. (Ball – 2004).

Social networks spread new ideas

Innovation was thought to follow a bell curve with a few early adopters, followed by the majority and then the laggards. Barabasi cites a 1954 study that showed that new medicines spread along social networks. It turned out that the early adopters and early majority were predominantly doctors with numerous social links. These highly connected doctors were more likely to be in touch with innovators, thus learning about the new drug more quickly. Once adopted by these doctors, the drug spread from these hubs to their less connected colleagues, who formed the late majority. Finally came the laggards, doctors who resisted adopting the new drug until the very end.[4] This highlights the importance of being linked into networks that spread innovation. These professional networks follow power law distributions.

The Internet links have a power law

Buchanan cites a study that reviewed 4,389 nodes in the network, linked together by 8,256 connections and found a power law pattern. (Refer Figure 36). Each time the number of Internet links doubles, the number of nodes with that many links becomes less by about five times. This simple power law relationship holds right across the board from nodes with only a few links to those with several hundred. The simplicity within the 'emergent' pattern suggests that however random and haphazard the Internet picture may look, it actually harbours a hidden order. The Internet is more or less pure hardware (i.e. hardware is the physical components of a computer system).[2]

---

[4]   Barabasi, Albert-Laszlo (2002). Linked – The New Science of Networks. Perseus Publishing. Cambridge, Massachusetts, USA. Pp. 113-129.

Figure 36
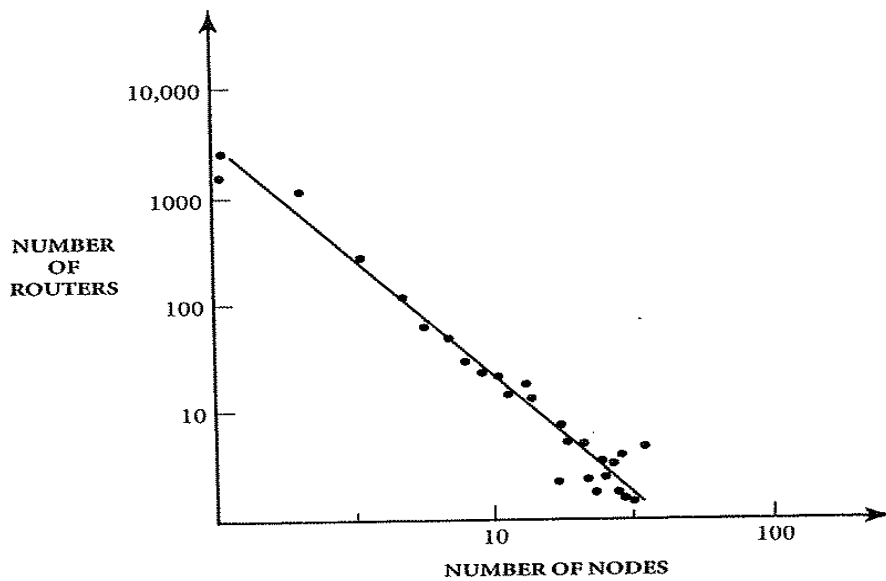
Power law curve for the Internet - The distribution of Internet 'nodes' according to how many links they possess. From Faloutsos et al. (Buchanan – 2002).

The World Wide Web links follow a power law

In contrast to the Internet, the World Wide Web is rather ethereal. This is the vast network of Web pages connected together by hypertext links. Barabasi and colleagues studied the Notre Dame Web site of 325,729 documents connected by 1,469,680 links and they also found a power law graph – identical to that for the Internet - the number of nodes having a certain number of links decreased by about five each time the number of links was doubled. (Refer Figure 37). In these networks, just a few nodes have so many links that 80 to 90 percent of the network's total number of links feed into just a small fraction of the nodes. So like all things that have a power-law pattern, there is a clear mathematical architecture of the log-log graph, an architecture that is dominated by especially well connected hubs.[2]

Figure 37
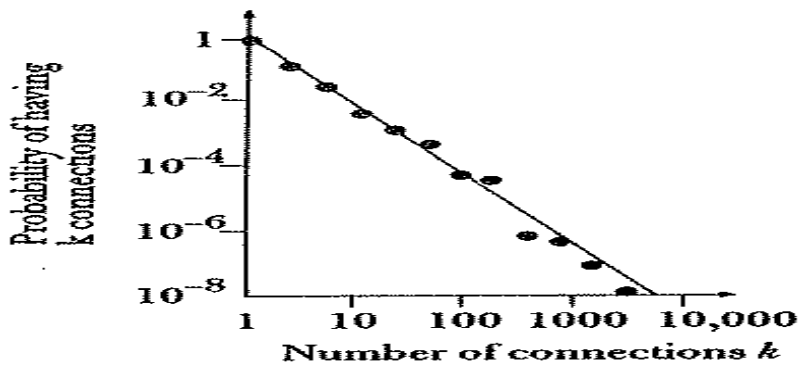
Power law curve for the World Wide Web (WWW) – The probability distribution of incoming hyperlinks to HTML documents on the WWW. The outgoing links have a similar distribution. (Ball - 2004).

107

In 1999 Albert Barabasi and colleagues found that the probability distributions of both incoming and outgoing links in this graph were power laws. That is the probability of a random page on the university (Notre Dame in Indiana) web site having a certain number of links to or from it depended on the number of links according to a power-law relationship. Most pages had few links; a few had many; and each time the number of links was doubled, the number of pages with that many links decreased by a constant factor. A power law is an emergent result as it is by no means the only relationship consistent with this expectation, i.e. it could have been a bell curve. But the power-law relationship says that there is no scale to the connectivity of the network. The system is scale-free. There is nothing pre-ordained in this power law. Everyone is completely at liberty to decide how many hyperlinks to make elsewhere. Yet out of this multiplicity of individual choices comes a mathematical law that holds true. The network seems to organise itself into this power law state, just as the sand-pile model of self-organised criticality generates a critical state with a power law distribution of avalanches.[3] The scale-free nature of the WWW link distribution indicates that collective phenomena play a previously unsuspected role in the development of the web, forcing us to look beyond the traditional random (normal distribution) graph models.[5] Because the World Wide Web is a 'small world', characterized by a power law and by scale free connectivity, anything can be found within a few clicks. It has been estimated that the best search engines reach only about 30 per cent of the entire indexable Web, and that some of them trawl through only 3 per cent.[3]

Resistance to attack

A cyber-terrorist who wants to trash the Internet need only identify a relatively small number of the most highly connected nodes and make those the targets of sabotage. In other words, when an attack is guided by intelligence it can do a disproportionate amount of harm to a scale-free network like the Internet. It is around the linchpin nodes that the most secure safety walls should be raised. Knocking out nodes at random will probably be futile, because the network can withstand a considerable amount of such punishment without losing much of its connectivity. Overall, the weakness of scale-free networks is their sensitivity to planned attack. If the hub nodes are removed, the network rapidly falls apart.[3]

Traffic jams follow a power law

Traffic jams also obey a power law. (Refer Figure 38). The kind of traffic jams referred to here are the ones on motorways where there is unobstructed dual carriageway, seemingly for no reason at all. Traffic flows smoothly when traffic is light, but once the amount of traffic on the road gets too great, traffic jams occur even though there seems to be no reason why everybody cannot move along smoothly. The problem is that if one vehicle brakes for any

5   Newman, Mark; Barabasi, Albert-Laszlo; Watts, Duncan J. (2006). *The Structure and Dynamics of Networks.* Princeton University Press. Princeton. New Jersey. U.S.A. Pp. 182.

reason (perhaps because the driver has simply got too close to the vehicle in front), there is a knock-on effect as everybody behind slows, and it is easier to slow down than it is to speed up.[6]

Both in computer modelling and in studies of real traffic flows we find a power law – the number of jams of different sizes (measured in terms of the number of vehicles involved) obeys the same kind of law as the number of earthquakes of different sizes. Just like real traffic jams, the jams in the model move backwards through the traffic, as more cars break behind the obstruction while those in front speed away, and just like real traffic jams they dissolve as the average speed of the traffic is slowed by the jam. As well as there being big jams and small jams, there are jams within jams, in a fractal pattern.[6]

Figure 38
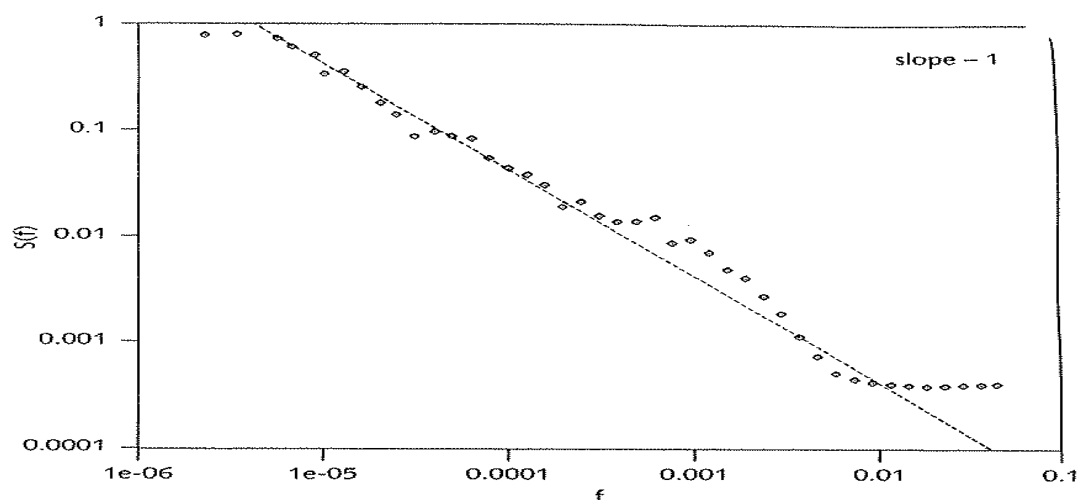
Power law curve for traffic jams, from Nagel et al. (Bak – 1997).

There are two practical lessons to be drawn. First, you don't need an unusually large 'trigger' (such as a crash) to cause a large jam; any size jam can be triggered by even the smallest disturbance; such as one car getting too close to the car in front and braking sharply. Nobody ever hits anybody, there is no obvious cause visible to most of the people involved (and the car that did cause the jam has probably sped off into the distance, its driver unaware of what is happening behind), but hundreds of people may be inconvenienced. Secondly, if the density of traffic increases, it can be kept flowing more freely by limiting its top speed, because this reduces the impact of the difference between the time taken to accelerate and the time taken to brake.[6]

The power-law graph (Figure 38) summarises in a mathematical form the irritating, unpredictable stop-start behaviour in traffic jams. Per Bak believes that the critical state, with

---

[6] Gribbin, John (2004). *Deep Simplicity. Chaos, Complexity and the Emergence of Life.* Allen Lane. An imprint of Penguin Books. London, England. Pp. 148-149.

jams of all sizes, is the most efficient state. The system has self organised to the critical state with the highest throughput of cars. If the density were slightly less, the road would be underutilised and if the density were slightly higher, there would be one big permanent, huge jam, absorbing a fraction of the cars. In both cases, the throughput would be less. The critical state is the most efficient state that can actually be reached dynamically.[7]

River networks follow a power law

One of the key discoveries of complexity is that finding lawlike patterns in complex nonequilibrium systems generally means taking a step back from the details, and focusing on the larger picture. We cannot predict accurately which person will be rich and which poor, or how large the next earthquake will be. The lawlike and predictable patterns emerge instead at the level of many events, in the statistics. Taking this step back from the prediction of events to prediction of general patterns is important if we wish to be able to answer different questions.[8]

At the detail level, the irregularity of rivers like the Mississippi is of little surprise because the specific layout of any river network reflects the geophysical history of the region, including earthquakes that shifted the river's course, patterns of rainfall, and so on. The overall structure of the river and its branches is like a fingerprint of that place and its history.[8]

When looking at river networks at a higher level, we can see that a power law applies. (Refer Figure 39). For example, the amount of water that any particular branch of a river network carries reflects the size of its 'catchment area', the total area that drains into it. The larger rivers, those carrying lots of water, are like the 'richer' rivers and the smaller rivers being the 'poorer'. Studies have found that the number of rivers with an amount W of water flowing through them (say per day) is proportional to $1/W^{\alpha}$, where α is about 1.43. Water flows through river segments in pretty much the same way that wealth accumulates in the hands of a few. A power law exists and they exist for trees and clouds also. This means that each time you double the amount of water drained, you find that the number of such streams falls by a factor of about 2.7. If there are one hundred streams draining a thousand square kilometres, then there will be only about thirty-seven draining two thousand square kilometres, and so on.[8]

This finding represents an impressive and rather unexpected simplicity for a network that seems to have little if any order (emergent property). This pattern applies to all river networks – the Nile, the Amazon, the Mississippi and the Volga etc. It points to a deep organizing

[7] Bak, Per (1997). *How Nature Works – The Science of Self-Organized Criticality*. Oxford University Press. Oxford, U.K. Pp. 196-198.

[8] Buchanan, Mark (2007). *The Social Atom – Why the Rich get Richer, Cheaters Get Caught, and Your Neighbour Usually Looks Like You*. Bloomsbury. New York. U.S.A. Pp. 174-177 and 228.

tendency behind the apparent disorder of all these networks. If river networks really were 'just' the product of a long string of historical accidents, with waters eroding channels here and there without regard for any overall plan, then this power law pattern would not exist. But the power law pattern suggests that there is a deeper reality and something else to discover behind these details. If the same pattern turns up in the Mississippi, the Nile, and the Amazon, details of geology and weather must have little to do with it.[9] (Refer Figure 39).

Because you don't need to take account of most physical factors (rocks, soil, rainfall, erosion etc), modelling river networks is far simpler than anyone would have guessed. All they needed to know was that each time the size of the basin doubled, the number of rivers fell by a factor of about 2.7. The modelled networks were never quite the same in terms of the detailed layout. Accidents and historical contingency play an important role here. History and chance are fully compatible with the existence of lawlike order and pattern.[9]

Scientists have found that the network of dry interlocking valleys on Mars looks like an archaic river network that dried out. It followed precisely the same power law as rivers on Earth.[8]

Figure 39
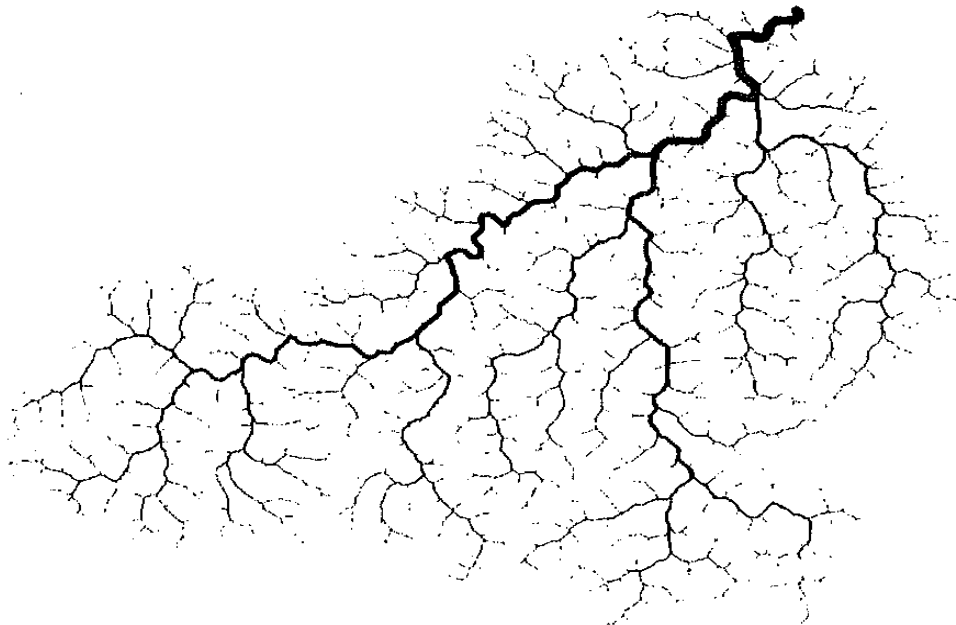
The structure of the Fella River network in northern Italy from Rodriguez et al. (Buchanan – 2002).

---

[9] Buchanan, Mark (2002), op. cit., Pp. 98-103.

The number of sexual partners follows a power law

A Swedish study (Liljeros) showed that the number of sexual partners also follows a power law, with most people having a few (up to 10 over a life time), some having dozens or more, and a few having hundreds. Our sexual relations network has a scale-free topology.[10]

Liljeros et al. (2001), carried out a sociometric study of sexual partnerships in Sweden. In the study 4781 individuals of ages 18 to 74 were polled, and, with a response rate of 59%, the experimenters obtained detailed information for 2810 nodes in the sexual contact network. The results indicated that the majority of respondents had between one and ten sexual partners during their lifetime, but a few had several hundred. It followed a power law – sexual contacts form a scale-free network – with exponent 3.5 for women and 3.3 for men (i.e. statistically the same). The outliers were the hubs in the network.[11]

Mark Newman and colleagues did their own study and confirmed the power law feature. They carried out a random digit dialling survey conducted among 1194 Seattle residents aged 18-39 years of age during 2003-2004. As Figure 40 shows, 38% of the survey reported five partners or fewer during their lifetime. However, the distribution also has a 'fat tail' representing a small fraction of the population who have a large number of partners.
15 percent of the population reported having 20 or more partners during their lifetime. Figure 40 is plotted on logarithmic scales and when viewed in this way the tail has an approximately straight-line form (dashed line), implying that the data follows a Pareto distribution or power law, at least approximately. Other studies have also made this observation. The spikes in the distribution are the result of digit preference, as respondents had a tendency to round off large numbers of partners to multiples of ten.[12] It is reasonable to assume that there is an exponential cut-off at the upper limits for the same reasons as networking of movie actors has an upper exponent cutoff. .

---

[10] Barabasi, Albert-Laszlo (2002), op. cit., Pp. 137.

[11] Newman, Mark; Barabasi, Albert-Laszlo; Watts, Duncan J. (2006). *The Structure and Dynamics of Networks.* Princeton University Press. Princeton. New Jersey. U.S.A. Pp. 180-181. Also, P 227-8 Fredrik Liljeros, Christofer R. Edling, Luis A. Nunes Amaralt, Eugene Stanley, Yvonne Aberg. The Web of Human Sexual connections. Nature, Volume 411, 21 June 2001. www.nature.com

[12] Foxman, Betsy; Newman, Mark; Percha, Bethany; Holmes, King K; Aral, Sevgi O. (2004?)

Measures of Sexual Partnerships: lengths, gaps, overlaps and Sexually Transmitted Infection. Pp1-23. University of Michigan. U.S.A.

Figure 40
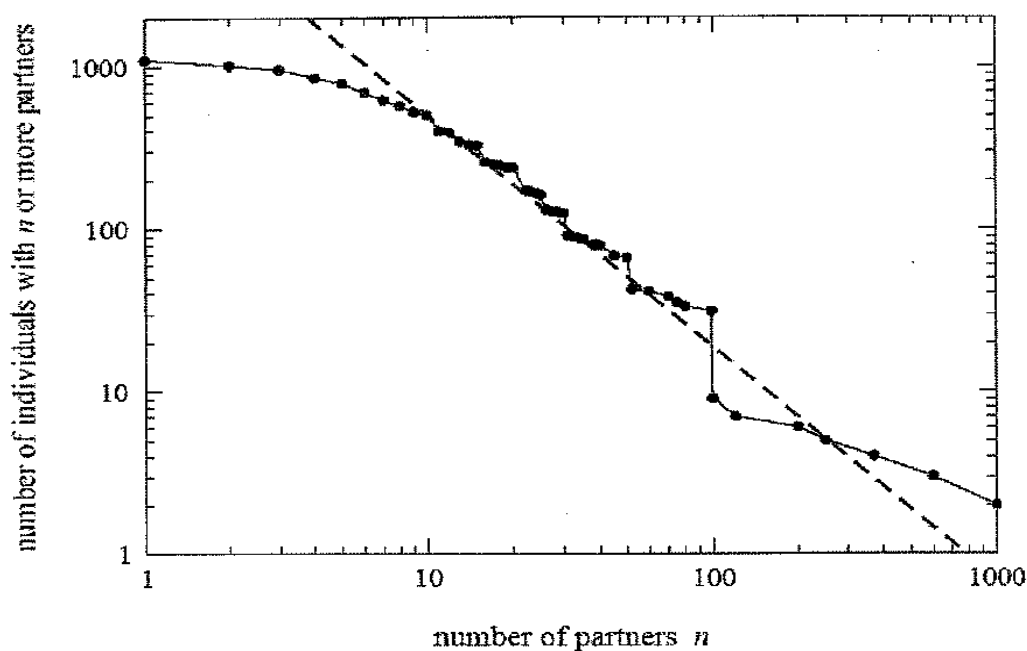
Power law curve for number of sexual partners. Cumulative distribution function of self-reported history of lifetime number of sex partners on logarithmic scales. The straight line in the tail of the distribution indicates a power law. The slope of the line is measured to be -1.4 (dashed line), which corresponds to a value of -2.4 for the exponent of the power law.[12](Foxman et al. – 2004?).

Power laws offer a way forward in dealing with venereal disease

One of the original, and still primary, reasons for studying networks is to understand the mechanisms by which diseases and other things (information, computer viruses, rumours) spread over them. For instance, the main reason for the study of networks of sexual contact is to help us understand and perhaps control the spread of sexually transmitted diseases.[13]

Public health programs ordinarily aim to vaccinate a certain fraction of the population more or less at random, hoping to drive the disease below the tipping point. If there is no such point, this will not work.[14] In traditional epidemiology, diseases display an epidemic threshold, as a function of their probability of transmission, between a regime in which only small local outbreaks of the disease occur and a regime in which the disease shows epidemic behaviour, reaching a significant fraction of the entire population in question.[11]

Epidemiologists conclude that the great majority of venereal disease cases could be accounted for by a small subset of the population – prostitutes, for one, but also servicemen stationed at a nearby military base and a small cadre of other men and women whose sexual activity was exceptionally high. Treating this core group is essential as this group is roughly

---

[13]   Newman, M.E.J. (2003), op. cit., Pp. 15 and 40.
[14]   Buchanan, Mark (2002), op. cit., P182-183.

113

identical to the connectors. On one hand, the connectors in a small-world network (i.e. a network that follows a power law) of the aristocrat kind (a name for the extreme power law outliers), are so well connected that their activity dominates all else, and no unfocussed treatment scheme has any chance of success. Well over 90 percent of all people could be immunized and the disease would keep right on going. On the other hand, treatment of only a fraction of the most highly connected offers a chance to bring the disease under control and possibly wipe it out. Actually applying this in the real world is not easy but it offers a basic game plan and a strategy that could help deal with Aids and for new diseases that might arise in the future.[14]

Network resilience is of particular importance in epidemiology, where 'removal' of vertices in a contact network might correspond for example to vaccination of individuals against a disease. Vaccination not only prevents the vaccinated individuals from catching the disease but also destroys paths between other individuals by which the disease might have spread.[13]

Pastor-Satorras and Vespignani (2001) demonstrated that networks with power law distributions may not have an epidemic threshold at all; in such networks it is possible for diseases always to be spread, regardless of the probability of transmission. They found that the cumulative distribution of the number of different sexual partners in one year decays as a scale-free power law that has a similar exponent for males and females. The scale-free nature of the web of human sexual contacts indicates that safe-sex campaigns are likely to be the most efficient way to prevent the spread of sexually transmitted diseases. These results are consistent with the preferential-attachment mechanism of scale-free networks: evidently, in sexual contact networks, as in other scale-free networks, 'the rich get richer'.[11]

The Liljeros team's most important finding is the scale free nature of the connectivity of an objectively defined, non-professional social network. This result indicates that the concept of the 'core group' considered in epidemiology studies must be arbitrary, because there is no well defined threshold or boundary that separates the core group from other individuals (as they would be for a bimodal distribution). Their results may have epidemiological implications, as epidemics arise and propagate much faster in scale-free power law networks than in single-scale normal distribution networks. Also, the measures adopted to contain or stop the propagation of disease in a network need to be radically different for scale-free networks. Single –scale networks are not susceptible to attack at even the most connected nodes, whereas scale-free networks are resilient to random failure but are highly susceptible to destruction of the best connected nodes. The possibility that the web of sexual contacts has a scale-free structure indicates that targeting of safe-sex education campaigns to those individuals with a large number of partners may significantly reduce the propagation of sexually transmitted diseases.[11]

Chapter 17 - Not all apparent power law distributions are really power law

Throughout my thesis I have accepted at face value all the claims by the numerous authors I have used as to which phenomena follow power law distributions and which don't. Since I am a student of strategic studies and not mathematics, I have only used examples that are described as power law distributions as these will clearly have the scale free properties that are so important in defining the difference between linear and non-linear phenomena. This distinction is what my thesis is about, since I believe power laws and the consequential punctuated equilibrium effects are important to many strategic questions (i.e. to all the non-linear ones).

This chapter considers a study by Mark Newman and his colleagues where they test twenty four data sets to confirm that they actually follow a power law distribution or whether they follow one of power laws' close statistical cousins such as log-normal or exponential distributions. They all apparently fit log-log graphs. For me, this is an interesting chapter as I am not sure that the mathematicians have worked out what the exceptions mean and so I assume that Newman and others are close to a knowledge frontier and those who rely on power law phenomena will need to continually monitor developments at this knowledge frontier. The importance for me of this chapter is to recognise that some apparent power law phenomena are not completely power laws. Since I had read Newman's articles before I started this thesis, I was therefore keen to include every graph of every subject covered, that I could, so that I could clearly see how different all these graphs actually looked in real life as opposed to reading about the effects in a narrative. This chapter effectively summarises Newman's conclusions and leaves out all the detailed mathematical reasoning.

Power-law distributions

Power laws are important for our understanding of non-linear phenomena and they appear in a diverse range of natural and man-made phenomena. For example, the populations of cities, the intensities of earthquakes, and the sizes of power outages are all considered power-law distributions. Quantities such as these are not well characterized by their typical or average values. For instance, according to the 2000 US Census, the average population of a city, town, or village in the United States is 8226. But this statement is not a useful one for most purposes because a significant fraction of the population lives in cities (New York, Los Angeles, etc) whose population is larger by several orders of magnitude.[1]

---

[1] Clauset, Aaron, Shalizi, Cosma Rohilla, and Newman M.E.J. *Power-Law Distributions in Empirical Data*. arXiv:0706.1062v2 (physics.data-an) 2 Feb 2009 www-personal.umich.edu/-mejn/pubs.html (Publications of Mark Newman) University of Michigan. Pp. 1-15.

The tool most often used for proving a power law distribution is to construct a histogram. The histogram is then plotted on double logarithmic axes. If in so doing one discovers a distribution that approximately falls on a straight line, then one can say that it may be a power law distribution, with a scaling parameter 'α' given by the absolute slope of the straight line. Typically this slope is extracted by performing a least-squares linear regression analysis on the histogram. This procedure dates back to Pareto's work on the distribution of wealth at the close of the 19th century. Newman believes that by itself this method has been found to be unsatisfactory, as systematic errors occur. Additional techniques are needed to determine whether distributions are or are not power law distributions.[1]

Problems with defining what a power law is

Newman believes that the detection and characterization of power laws is complicated by the large fluctuations that occur in the tail of the distribution, the part of the distribution representing large but rare events and by the difficulty of identifying the range over which power-law behaviour holds. Newman found that commonly used methods for analyzing power-law data, such as least-squares fitting, can produce substantially inaccurate estimates of parameters for power-law distributions. Newman's approach combines maximum-likelihood fitting methods with goodness-of-fit tests based on the Kolmogorov-Smirnov statistic and likelihood ratios.[1]

Most empirical studies of ostensibly power-law distributed data have not attempted to test the power-law hypothesis quantitatively. Instead, they typically rely on qualitative appraisals of the data, based for instance on visualization. But these can be deceptive and can lead to claims of power-law behaviour that do not hold up under closer scrutiny. Two other patterns are also possible, being a log-normal distribution and an exponential distribution. Because each of these distributions looks roughly straight line on the log-log plot, one might, upon a cursory inspection, judge all three to follow power laws, albeit with different scaling parameters. This judgement would, however, be wrong – being roughly straight on a log-log plot is a necessary but not sufficient condition for power-law behaviour. Newman does not provide any simple guidance on what would be a better definition of power law distributions.

Because in all observed distributions there are small deviations because of the random nature of the sampling process, there will always be a challenge determining which power-law distributions are and which are not. There are practical difficulties with test data. For example, very small data sets should always be treated with caution, since it is difficult to rule out alternative fits to such data, even when they are truly power-law distributed, and conversely the power-law form may appear to be a good fit even when the data are drawn from a non-power-law distribution.[1]

The Clauset/ Newman power law quantitative study

Newman tested twenty-four data sets drawn from a broad variety of different branches of human endeavour, including physics, earth sciences, biology, ecology, paleontology, computer and information sciences, engineering, and the social sciences.[2]

They are as follows:

A.   The frequency of occurrence of words in the novel Moby Dick by Herman Melville.

B.   The degrees (i.e., numbers of distinct interaction partners) of proteins in the partially known protein-interaction network of the yeast Saccharomyces cerevisiae.

C.   The degrees of metabolites in the metabolic network of the bacterium Escherichia coli.

D.   The degrees of nodes in the partially known network representation of the Internet at the level of autonomous systems for May 2006. An autonomous system is a group of IP addresses on the Internet among which routing is handled internally or 'autonomously', rather than using the Internet's large scale border gateway protocol routing mechanism.

E.   The number of calls received by customers of AT&T's long distance telephone service in the United States during a single day.

F.   The intensity of wars from 1816-1980 measured as the number of battle deaths per 10,000 of the combined populations of the warring nations.

G.   The severity of terrorist attacks worldwide from February 1968 to June 2006, measured as the number of deaths directly resulting.

H.   The number of bytes of data received as the result of individual web (HTTP) requests from computer users at a large research laboratory during a 24-hour period in June 1996. Roughly speaking this distribution represents the size distribution of web files transmitted over the Internet.

I.   The number of species per genus of mammals. This data set, compiled by Smith et al., is composed primarily of species alive today but also includes some recently extinct species, where 'recent' in this context means the last few tens of thousands of years.

J.   The number of sightings of birds of different species in the North American Breeding Bird Survey for 2003.

K.   The number of customers affected in electrical blackouts in the United States between 1984 and 2002.

---

[2]   Ibid., Pp. 22-30.

L. The number of copies of bestselling books sold in the United States during the period 1895 to 1965.

M. The human populations of US cities in the 2000 US census.

N. The size of email address books of computer users at a large university.

O. The sizes in acres of wildfires occurring on Federal land between 1986 and 1996.

P. Peak gamma ray intensity of solar flares between 1980 and 1989.

Q. The intensities of earthquakes occurring in California between 1910 and 1992, measured as the maximum amplitude of motion during the quake.

R. The number of adherents of religious denominations, bodies, and sects, as compiled and published on the web site 'adherents.com'.

S. The frequencies of occurrence of US family names in the 1990 US Census.

T. The aggregate net worth in US dollars of the richest individuals in the United States in October 2003.

U. The number of citations received between publication and June 1997 by scientific papers published in 1981 and listed in the science Citation Index.

V. The number of academic papers authored or co-authored by mathematicians listed in the American Mathematical Society's MathSciNet database. Data compiled by J. Grossman.

W. The number of 'hits' received by web sites from customers of the America Online Internet service in a single day.

X. The number of links to web sites found in a 1997 web crawl of about 200 million web pages.

These data sets are shown in Figures 41 and 42. Note the range of subjects covered – from religions to surnames.

Many of these data sets are only subsets of much larger entities (such as web sites, which are only a small fraction of the entire web). In some cases it is known that the sampling procedure used to obtain these subsets may be biased, as, for example, in the protein interactions, citations and authorship, and the Internet. Newman did not attempted to correct any biases in their analysis.[2]

Newman found that the testing tools, such as likelihood ratio tests do not give conclusive answers. Many of the twenty four phenomena are plausible power laws, but they are also plausible log-normal's and stretched exponentials. In general, it is extremely difficult to tell the difference between log-normal, exponential and power-law behaviour.[2]

Figure 41
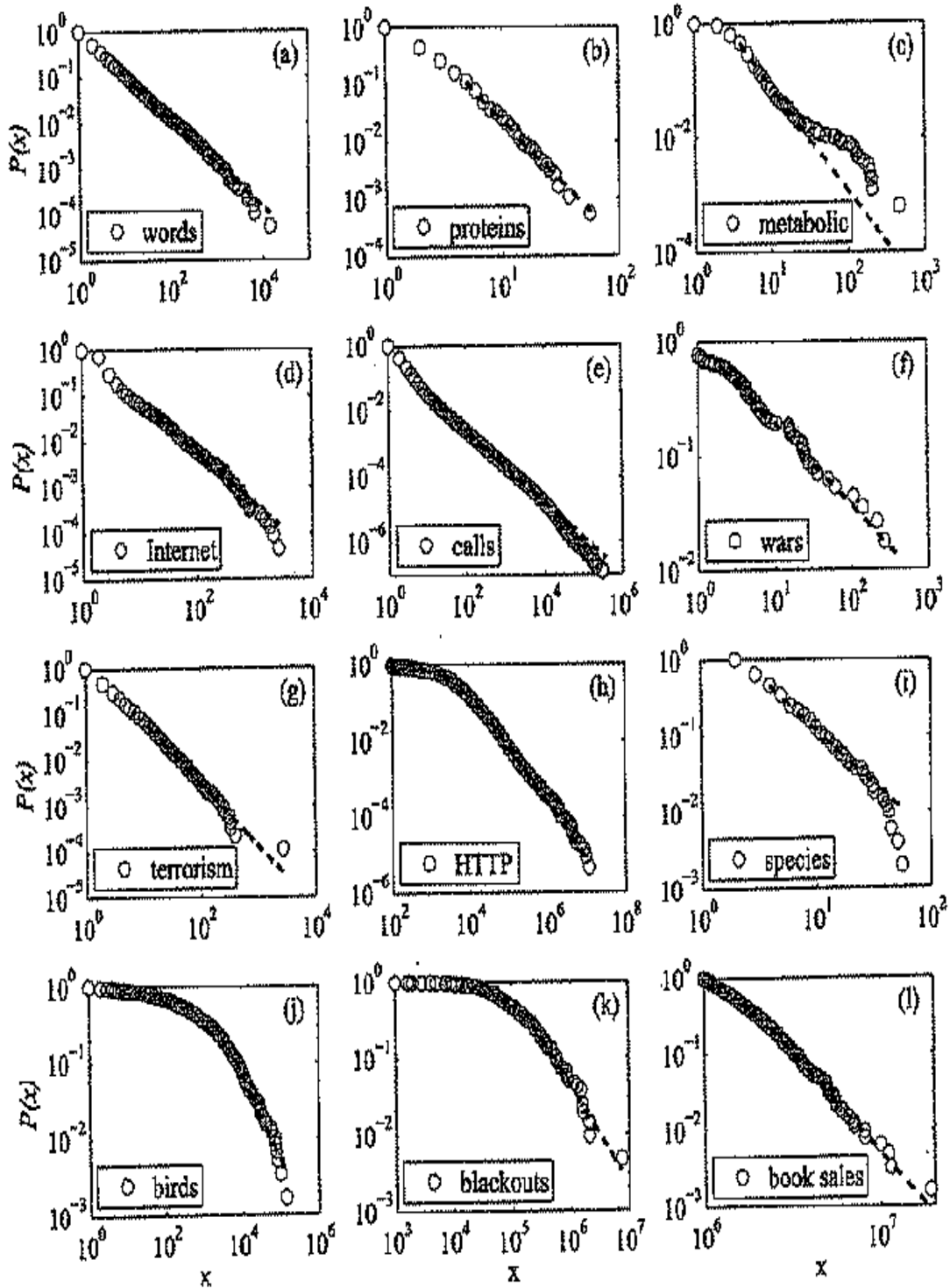
Testing for power-law fit – Samples (a) to (l). (Clauset, Shalizi, Newman 2009).

119

Figure 42
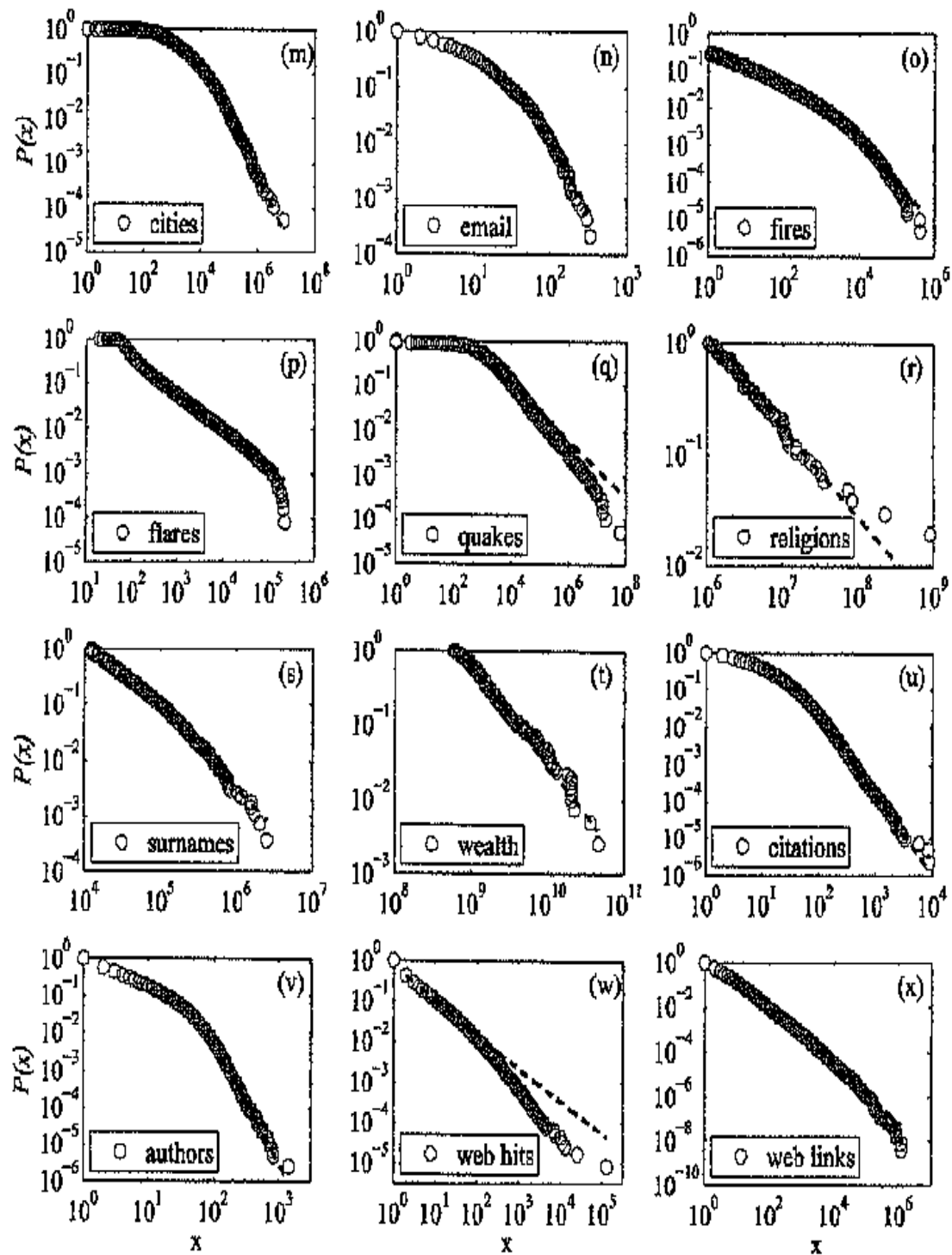
Testing for power-law fit – Samples (m) to (x). (Clauset, Shalizi, Newman 2009).

Summary of 'Support' for a power law distribution

Power-law distributions come in two basic types: continuous distributions concerning continuous real numbers and discrete distributions where the quantity of interest can take only a discrete set of values, typically positive integers. Both sets of data need to be treated differently.[1] Following, are two tables (Tables 5 and 6) with the different data sets sorted into

continuous and discrete. The tables are a summary of the Newman results with the final column of each table showing the Newman team's judgement of the statistical support for the power law hypothesis for that data set. They have used the words 'None', 'Moderate', 'Good' and 'With cut-off' to describe their judgements. These terms are defined below.

Table 5

Data set - Continuous (non-discrete) data	Sample size n	Power law exponential α	Support for power law
Birds species sightings	591	2.1(2)	moderate
Blackouts (x10³)	211	2.3(3)	moderate
Book sales (x10³)	633	3.7(3)	moderate
Cities (x10³)	19 447	2.37(8)	moderate
Forest Fire size (acres)	203 785	2.2(3)	With cut-off
Flares, solar intensity	12 773	1.79(2)	With cut-off
HTTP size (Kilobytes)	226 386	2.48(5)	none
Earthquakes, intensity (x10³)	19 302	1.64(4)	With cut-off
Religions, followers (x10⁶)	103	1.8(1)	moderate
Surnames, frequency (x10³)	2753	2.5(2)	With cut-off
Wars, intensity	115	1.7(2)	moderate
Wealth. Net worth (mil. USD)	400	2.3(1)	none
Web hits	119 724	1.81(8)	With cut-off
Web links	241 428 853	2.336(9)	With cut-off

Table 6

Data set – Discrete (integer)	Sample size n	Power law exponential α	Support for power law
Internet degree	22 688	2.12(9)	With cut-off
Calls, telephone received	51 360 423	2.09(1)	With cut-off
Citations to papers	415 229	3.16(6)	moderate
Email address books size	4581	3.5(6)	With cut-off
Metabolic degree	1641	2.8(1)	none
Papers authored	401 445	4.3(1)	moderate
Protein interaction degree	1846	3.1(3)	moderate
Species per genius	509	2.4(2)	With cut-off
Terrorist attack severity	9101	2.4(2)	moderate
Words, use	18 855	1.95(2)	good

The final column of the table lists their judgement of the statistical support for the power law hypothesis for each data set. These are:

'**None**' indicates data sets that are probably not power-law distributed.

'**Moderate**' indicates that the power-law is a good fit but that there are other plausible alternatives as well.

'**Good**' indicates that the power-law is a good fit and that none of the alternatives considered is plausible.

'**With cut-off**' means that the power law with exponential cut-off is clearly favoured over the pure power law. In each of these cases, however, some of the alternative distributions are also good fits, such as the log-normal or the stretched exponential distribution.[2]

No standard and easy method of analysis to determine power laws

The study of power laws is relatively new and the methodology for this important subject is still being developed. At this stage there are no easy or standard methods for analyzing power-law data. As Newman points out, this naturally leaves open the possibility that apparent power-law behaviour is, in some cases at least, the result of wishful thinking.[2]

The common practice of identifying and quantifying power-law distributions by the approximately straight-line behaviour of a histogram on a doubly logarithmic plot should not be trusted completely. Such straight-line behaviour is a necessary but by no means sufficient condition for true power-law behaviour.[2]

Newman cautions that for many of these data sets the power-law hypothesis turns out to be, statistically speaking, a reasonable description of the data. That is, the data are compatible with the hypothesis that they are drawn from a power-law distribution, although they are often compatible with other distributions as well, such as the log-normal or stretched exponential distributions. In the remaining cases the power-law hypothesis is found to be incompatible with the observed data. In some instances, such as the distribution of earthquakes, the power law is plausible only if one assumes an exponential cut off that modifies the extreme tail of the distribution.[2] Newman believes that the data sets that are considered problematic, indicate complex underlying processes that merit further study.[1]

Table 5 shows that wealth (Refer Pareto's law – Chapter 5) is deemed by Newman not to be a pure power law. It is unclear what that inherently means as I don't understand enough mathematics to know what it means for non-linear log-log graphs not to be a power law. They will still be non-linear, so we need to wait until the body of knowledge on this subject addresses this issue and explains what it means. For the purposes of my study, Pareto's law of wealth is still a very useful explanation of a non-linear distribution and is widely quoted (perhaps mis-quoted) in complexity literature. Newman's 2009 study highlights that in any developing area, initial classifications may have to be modified.

Chapter 18 – Strategic implications of non-linearity

For me, this thesis has been an important stepping stone in my study of strategic thinking, as it built on my STRA 519 Research Essay which also considered forecasting. The STRA 519 essay considered linear approaches which by the time I had finished that essay, I considered that I had mastered (i.e. had a good understanding of), whereas from my learning point of view, this thesis covers non-linearity (complexity) and especially the study of power laws that is almost completely new territory for me and it has opened up completely new possibilities. These possibilities are outlined below.

Implications of complexity when considering strategic forecasting

One way I know that I have advanced my understanding of strategic forecasting from doing this thesis, is from the following example. Recently, after I had scripted most of this thesis, I sat down and read the following future focused books:

- *A Brief History of the Future* (2009) by Jacques Attali,
- *The Next 100 Years – A Forecast for the 21st Century* (2009) by George Friedman,
- *The World is Flat – A Brief history of the twenty-first century* (2005) by Thomas L. Friedman,
- *Hot, Flat, and Crowded – Why we need a green revolution and how it can renew America* (2008) by Thomas L. Friedman,
- *The Decline of American Power – The U.S. in a chaotic World* (2003) by Immanuel Wallerstein, and
- *The Post American World* (2008) by Fareed Zakaria.

I realised that all of these books gave me good safe 'linear' advice if I was advising the Mayor of the Auckland Council or Prime Minister or some other decision maker. For example, they all talk about the relative decline of American power and the relative and absolute rise of many emerging powers (China, Russia, India etc). Most developed countries are aging and will be short of workers etc. These books are authoritive and full of important and valid trends, but I suggest the phrase repeated by Thomas Friedman, 'Trend is not destiny', is a critical point. All of the books above use standard extrapolation scenario techniques based on trends or 'mega trends' which are mostly fairly linear. They don't factor in non-linear turbulence and sudden, dramatic, transformational and unexpected changes to the established order. It was only when I read the following strategy book that I saw for the first time someone linking the 'sandpile' criticality of Per Bak to political and strategic events in the world.

- *The Age of the unthinkable – Why the new world disorder constantly surprises us and what we can do about it* (2009) by Joshua Cooper Ramo.

The following paragraphs give a flavour of a non-linear approach to strategic forecasting from Joshua Ramo (Kissinger Associates)[1]. All the examples are far removed from linear-equilibrium thinking.

a)   Increasing the connectedness, increases the future uncertainty

Ramo believes our future is more dangerous than we think. There exists a dangerous paradox – the more closely we're bound, the less resilient we all become. Studies of food webs or trade networks, electrical systems and stock markets, find that as they become more densely linked they also become less resilient; networks, afterall, propagate and even amplify disturbances. Worse, the more efficient these networks are, the faster they spread those dangers. Drop a shock into a network and you get a chain reaction and the end state can not be predicted.[2]

b)   Politics and foreign affairs tend to move towards the 'critical' state

When Per Bak described the 'tendency of large systems to evolve into a poised 'critical' state, way out of balance, where minor disturbances may lead to events, called avalanches, of any size', he could have been speaking about the Middle East, relations between the United States and China, the oil market, disease, nuclear proliferation, cyber warfare or a dozen other problems of global affairs and security. We are organised for instability. Bak believed that sandpile energy, the energy of systems constantly poised on the edge of unpredictable change, was one of the fundamental forces of nature. Bak's sandpile universe was violent – and history making. It wasn't that he didn't see stability in the world, but he saw stability as a passing phase, as a pause in a system of incredible and unmappable non-linear dynamism. This dynamic sandpile energy demands that we accept the basic unpredictability of the global order.[2]

c)   Top down 'Grand Strategies' don't work in a dynamic complex world

Ramo believes that the sandpile effect takes notions at the heart of contemporary foreign policy; ideas like soft power or that democracies won't fight one another, and shreds them. Traditional Grand Strategy took strength from its static nature. It began with a vision of the world as we wanted it to be – all democratic or all French or all British etc. Old grand strategies looked like the Cold War détente balance between East and West or the Monroe Doctrine. They relied on the idea of rational powers, threats that could be named and monitored, violence that came in the form of armies, and usually clear lines between economic, environmental, political, and military policy. In the sandpile world no such division is possible. In a world that is changing fast, we need a grand strategy that's capable of the sorts of rapid change the world itself produces, because much of what we have to confront

---

[1]   Ramo, Joshua Cooper (2009). The Age of the unthinkable – Why the new world disorder constantly surprises us and what we can do about it. Little, Brown. London U.K. Pp. 198.

[2]   Ibid., Pp. 49-60.

will be things that have never occurred to us before. Ramo believes we don't need a grand-strategy, but we do need 'deep security', that assumes that we don't know all the answers or have anticipated all possible events. We need to be as adaptive and flexible as the world is.[3]

d)   The unexpected fall of the Soviet Union

The end of the Soviet Union was one of those sandpile avalanches that, like the 1929 stock market collapse, demanded a complete remapping of the world. The end of the U.S.S.R. stunned the same minds that were later stunned by 9/11 and the financial collapse of 2008. The U.S.S.R. imploded and was seen as a sign of the triumph of U.S. ideals – for example, the inherent supremacy of democracy and capitalism over other systems. This was the 'linear' – 'input Western ideas, output implosion' explanation.[4]

The Russians including Gorbachev (who Ramo interviewed) were baffled by the collapse of the U.S.S.R. The Soviet Union had been exposed to periods of far greater economic and social pressure without imploding. The U.S. military threat had been part of Soviet life for decades. If the political system was so deeply fragile and rotten, then why had the collective predictive superstructure of Soviet and foreign experts missed the signs? [5]

The Soviet Union was actually one of those critically poised systems.   As Per Bak said, 'Small things can trigger a big change, so you don't need a big cause for a big event'. The U.S.S.R. collapsed quickly. Kotz and Weir's research concluded that the U.S.S.R. didn't inherently collapse because of popular pressure upwards from the grass roots of Soviet life – pressure that could have been suppressed at any time – but largely because of the 'nomenklatura'. Politburo members like Gorbachev certainly held sway on the macro level, but the nomenklatura (e.g. army officers, professors and officials etc) who actually managed the day to day workings of the U.S.S.R., decided once Gorbachev began reforming a system that had protected their rights and privileges, they had more to gain by letting the U.S.S.R. fracture than by holding it together. The nomenklatura logic was that when the empire fell, they would be in the best position to pick up the pieces. This was perhaps a cold, selfish decision. Weir and Kotz concluded that the surprisingly peaceful and sudden demise of the U.S.S.R. happened because it was abandoned by most of its elite. It was not a real revolution like 1917 with a new set of masters, but one where the new leaders like Putin, were also previously elite members. It was not a clear cut and inevitable Western 'triumph' in the Cold War. Gorbachev told Ramo it is wrong to assume that the West won the Cold War. Change produces unpredictability and surprise. Gorbachev made his system more complex and unstable rather than less.[5]

---

[3]   Ibid., Pp. 107-8.

[4]   Ibid., Pp. 65-72.

e)   A small action can lead to big effects (non-linear effect)

In 1982 the British were desperate to stop the flow of Argentine soldiers arriving at the lone airport at Stanley, in the Falklands. There was only one plane that could get there, which was the Vulcan bomber and it needed multiple mid-flight refueling to get to the Falklands and back. It was very high risk. The question was raised as to how many bombs were needed to take out the airfield. A Defence scientist (James Moffat) worked this out, prepared a map, and sent it to his seniors. They passed it within hours to Prime Minister Thatcher who immediately approved the Vulcan planes to carry out the difficult mission. As Moffat had predicted, one bomb had hit the runway hard enough to knock it out. Suddenly and quite unexpectedly the Argentine Airforce began shifting their planes back to Buenos Aires in a defensive move. The bombings had shown the Argentines that the RAF had the potential to hit Argentina and to the Argentines that was deeply disturbing, especially to the generals who now felt personally threatened. It wasn't that Moffat had managed to take out the Stanley airport; he had indirectly taken out a large chunk of the entire Argentine air force. A small action had a big unexpected impact. On reflection it was as if one bomb, exactly placed, is as powerful as an entire Airforce.[5]

f)   Creating what you fear most - your own complex adaptive enemy

Ramo believes that Hizb'allah (Pro-Palestinian resistance/ terrorist movement) has a capacity for creativity and innovation, due to the need to survive under the full pressure of Israeli attack. Ramo believes it is an example of a complex adaptive system in operation. The Israelis over time have forced their enemies to be more innovative and to evolve into effective networks. Those that didn't were put out of action. Those that evolved grew stronger. As a condition of its survival during the years of its fight against Israel, Hizb'allah has become a machine for innovation, as much terror lab as terror group.[7]

In 2006 the Israelis used powerful air strikes to pound the visible Hizb'allah to kill the invisible. To Ramo, this is classic old-school thinking. However, unfortunately for the Israelis, fewer than 500 Hizb'allah fighters managed to frustrate a 30,000-man Israeli attack, including one of the most extensive air campaigns in Middle East history. Hizb'allah, to prove their point, had made a show of firing the same number of missiles on the last day of the war as they had on the first.[6]

This created a dangerous problem for Israel; direct attacks on Hizb'allah made the militants more resilient, not less. Over time Israel's small engagements meant Hizb'allah's resilience improved as they were able to shift and learn and change – and they did all of those things even better when they were under attack. The Israelis have not only failed to eliminate the Hizb'allah, they have made them stronger.[7] To be fair to the Israelis, they can also learn and

---

5   Ibid., Pp. 200-203.

6   Ibid., Pp. 4-9, 141-2, 183-190.

adapt to new situations. Recent computer attacks on Iran's centrifuge facilities at their power plants shows that everything is fluid in conflicts/ warfare and learning is essential.

g)   Success of networked and decentralized military leadership

Military historians have studied the bewildering efficiency of armies that swarm like bees or ants, highly decentralized groups that bend, adjust, and attack based on a far better sense of local conditions than any central commander could ever have. This form of warfare flew in the face of centuries of command-and-control warfare theory, but that is what made the German U-boats so effective, the Mongol horsemen, Mao Zedong, the Hizb'allah and others so successful. Swarming is the classic immune-system response. This kind of self-organisation exists in many of the most efficient and resilient systems in the world.[7]

The important point from a strategy perspective is to acknowledge the success of bottom up, decentralized, non linear approaches and to equally acknowledge the success of top down, centralized, and linear approaches. Each approach suits different circumstances and both can apply to different aspects at the same time. For example in most of the examples above, such as the German U-boats or Mao Zedong, there was clearly a strong central direction and leadership even if there was some flexibility of methods and timing at a local level. The linear/ non-linear debate is not saying that complexity is the answer for all situations, only that non-linear solutions are the most appropriate for non-linear situations, and linear solutions are best for linear situations. Historically all situations whether linear or non-linear, were treated in a linear top down approach. The reality for all networked organisations is to apply the newly developing non-linear complexity approaches, which may turn out to be for most situations.

Whilst all the examples that I have just quoted from Ramo are quite dramatic, they need to be put into a coherent strategic framework. The next section gives one such framework.

The changing face of strategic planning due to non-linearity

Markku Sotarauta[8], who teaches strategic thinking for regional economic development in Finland, provides the following table (Table 7) which effectively summarises the linear/ non-linear divide. Sotarauta provides a good high level summary of my summary of key words used to compare linear and non-linear effects of Table 2 in Chapter 2; as well as covering many of the issues described by Ramo above. What Sotarauta calls Governance of Ambiguity is the approach that acknowledges the world has non-linear aspects with power-law networks that have tipping points, cascades or avalanches. It also acknowledges that combinations of approaches are needed so that a strategy may have aspects that are top down as well as bottom up and that like most New Zealand public sector strategies they are developed and

---

7   Ibid., Pp. 236-7.
8   Dator, James A. (2002). Advancing futures - Futures Studies in Higher Education. Chapter 16, Teaching Futures-Seeking Policy process by Markku Sotarauta. Praeger Studies on the 21st Century. Westport, Connecticut. United States of America. Pp. 217 to 231.

implemented by multiple parties, with multiple viewpoints and the whole process has a 'herding feel' to it. This contrasts with the historical view of strategy which is perhaps more like the Biblical 'Moses' approach, where the leader develops the plan and then orders the hierarchy they control to implement it. Sotarauta uses the tiles of each approach to reflect their underlying thinking. 'Government' is the state, such as the Crown or Ministries or other public sector organisations. 'Governance' includes the 'Government' but also includes other key stakeholders such as the education sector, Non-Governmental Organisations (N.G.O.s) and the business sector. Governance is therefore a much more inclusive approach to strategy development. 'Uncertainty' has a focus on risk reduction, whereas governance of 'ambiguity' is accepting a greatly increased sense of uncertainty for most aspects of the strategy development.

Table 7 Characteristics of Government of Uncertainty and Governance of Ambiguity

	Government of Uncertainty (Historical approach)	Governance of Ambiguity (Model from the 1990s onwards)
Development view	• Linearity, predictability, and continuity • Shared worldviews and conceptions of future • Cause-effect relations	• Evolutionary view of development; complexity, discontinuity, unpredictability • Different worldviews and diverging conceptions of future • Cause-effect relations blurred over time
Structure	• One dimensional • Centralized integration • Hierarchies • Division into units	• Multidimensional • Decentralized integration • Networks • Entities and parts
Governance processes	• Top-down • Unilateral dependence • Strong, do-it-alone dependence principle • State dependent	• Top-down and bottom-up • Multidimensional • Do-it-together principle emphasized • Dependent internationally, nationally, regionally, and locally
Planning	• Formal and hierarchical • Separation of planning and implementation • Approaches and analyses based on averages, from rules to exceptions	• Network management, communicative planning • Planning, decision making, and implementation differences blurred • Approaches and analyses for each situation, from exceptions to rules.

So far in this chapter I have described various examples where the strategic reality is non-linear and I have outlined one framework that essentially shows the historical approach which had a linear/ top down focus and the modern approach of acknowledging non-linear aspects. This change over time may seem very straight forward, but I believe there is a definite 'edge' to why acknowledging and understanding that non-linear thinking in strategy is essential for

our very survival and that is not a trivial point. For example, all the major problems facing humanity, such as climate change or obesity, are called 'wicked complex problems' and they can only be solved by using non-linear approaches that understand such things as power-law cascades and feedback systems etc. The next section considers wicked problems and why the solutions to them have everything to do with non-linear approaches.

Tackling wicked problem

The term 'wicked' was first used in 1973 by two Americans - Rittell and Webber[9] who noted that there was a whole realm of social planning problems that cannot be successfully treated with traditional linear, analytical approaches. They called these issues wicked problems and contrasted them with 'tame' problems. Tame problems are problems that can be tightly defined and a solution fairly readily identified or worked through. Linear approaches could not handle the social complexity, conflicting policy objectives and approaches, multiple organisations and stakeholders, unintended consequences and imperfect knowledge of what are now called wicked problems. The handling of wicked problems requires holistic rather than linear thinking.

Wicked problems have the following attributes: The problems are difficult to clearly define, they have many interdependencies and are often multi-causal, actions lead to unforeseen consequences, the problems are not stable, there is usually no clear solution, they are socially complex, they can not be solved by one agency, they involve changing behaviour and they are characterised by chronic policy failure.[9]

The Australian Public Service Commission[9] suggests that tackling wicked problems requires at least the following:

- Holistic and big picture and not partial or linear thinking
- Innovative, adaptable and flexible approaches
- The ability to work across agency boundaries
- Increasing understanding and stimulating a debate on the application of the accountability framework
- Effectively engaging stakeholders and citizens in understand the problem and in identifying possible solutions
- Additional core skills to just high-level analytical, conceptual, writing and project management; to include communications, influencing skills and the ability to work co-operatively.
- A better understanding of behavioural change by policy makers
- A comprehensive focus and/or strategy

---

[9] Australian Government/ Australian Public Service Commission (2007). *Tackling Wicked Problems – A Policy Perspective*. Contemporary Government Challenges. http://apsc.gov.au/publications07/wickedproblems.pdf

- Tolerating uncertainty and accepting the need for a long-term focus.

These solutions clearly tie back to the Table 7 analysis. We live in a world of very complex 'wicked problem' of which any number of them could threaten our survival. It is paramount to understand and apply the best strategic tools we have for solving these problems and understanding non-linearity and systems based thinking are important aspects of this.

Chapter 19 – Conclusions

Forecasting, when power law distributions apply

Power law statistical distributions can be seen in a wide variety of natural and man-made phenomena, from earthquakes and solar flares to populations of cities and sales of books. This sheer diversity of effects that have power law distributions is an important fact that has only become evident over the last decade or so. However, it is important to remember that few of these real world distributions follow a power law over their entire range, and in particular not for smaller values of the variable being measured. That is why my thesis includes when ever possible the actual power law graphs, so that the imperfect reality can be seen. A second important fact is that the power law effects are usually emergent properties and would never have been detected by linear approaches and explanations. Overall, complexity has its own set of terms and methods and some of the empirical sample sizes for the graphs are impressive.

The answer to my research question, 'Forecasting, when power law distributions apply', is to understand the 'sandpile' metaphor of Per Bak, where the next event could be very small or occasionally very large, and that the overall pattern of the events will follow an emergent power law 'log-log' chart. Under these conditions forecasting what the next event will be becomes very problematic, if not downright impossible. Patterns can be predicted but not point events. Big events do not require big causes. Trends are not destiny.

Throughout the thesis the main ideas and concepts that strategic forecasting needs to apply to non-linear power-law distributions are:

- Understanding the importance of complex adaptive systems,
- 'Scale-free' or 'scale invariant' features,
- Punctuated equilibrium (periods of stability and periods of rapid change),
- Self-organised criticality (poised at a tipping point),
- The 'Butterfly' effect (small changes can have big effects),
- Emergent properties or patterns,
- The 'Matthew' effect (winner takes all),
- Change is caused by both internal/ external triggers,
- Pattern prediction rather than point prediction,
- Straight line log-log graphs are the signature of power law-distributions,
- With emergent power-law patterns, details do not matter,
- Increased networks/ interdependence, means increased turbulence, and
- Strategy is about surviving the 'Red Queen' Race.

These concepts only apply to non-linear power-law distributions and they need to be understood in that uniquely non-linear way, where the whole is greater than the sum of the parts. Each of these concepts is considered in greater detail below.

Power law distributions only apply to 'complex adaptive systems'

Nonlinear systems are systems in which changes in initial conditions do not produce proportional changes in the final state.[1] Non-linear systems are common, of which complex adaptive systems are an important subset. Complex adaptive systems involve a medium-sized number of agents (drivers, traders, molecules etc) who are generally both intelligent and adaptive. These agents make decisions in accordance with various rules, and are ready to modify their rules of action on the basis of new information that comes their way. There are no dictators or centralised controllers in these systems. There is generally no single driver, trader, or molecule that has access to what other actors in the system are doing, so the agents in a complex adaptive system make their decisions and update their rules for action on the basis of local rather than global, information.[2]

Power law distributions are 'scale-free' or 'scale invariant'

Power law distributions decay far more slowly than the normal, Bell or Gaussian curve, allowing for large 'rare events' that can never happen in a normal curve. These large 'rare' events are an important characteristic of power law distributions. Power laws formulate in mathematical terms the notion that a few large events carry most of the action.[3]

The power-law distribution is the only distribution that is scale-free. This contrasts with what happens under most normal Gaussian linear distributions which are inherently stable and do have a 'scale' or mean. In Gaussian systems there are no dramatic tipping points or wild fluctuations. In a Gaussian/ normal distribution, there is an average/ mean and practically all the values are tightly clustered (within three standard deviations) around that figure. In a power law distribution there is no valid average figure that represents the majority of values. The power law average combines the numerous low figures with a few extreme figures and the answer is much greater than the median because of the high outliers. This forces us to abandon the idea of an average. There is no intrinsic scale. This is the reason power-law distributions are called 'scale-free'.[4]

[1] Thuan, Trinh Xuan (2001). *Chaos and Harmony. Perspectives on Scientific Revolutions of the Twentieth Century.* Oxford University Press. New York. USA. Pp. 340-346.
[2] Casti, John L (1997). *Would-Be Worlds. How Simulation is Changing the frontiers of Science.* John Wiley and Sons, Inc. New York. U.S.A. Pp. Ix-x.
[3] Barabasi, Albert-Laszlo (2002). Linked – The New Science of Networks. Perseus Publishing. Cambridge, Massachusetts, USA. Pp. 67-72.
[4] Newman, M.E.J (2006). *Power laws, Pareto distributions and Zipf's law*. arXiv:cond-mat/0412004v3 (cond-mat.stat-mech) 29 May 2006.

For example, there is no 'typical' size for a city in the USA or elsewhere, and no reason to see special historical or geographical situations behind the emergence of the very biggest. Likewise there is no typical size computer company. Do you include Microsoft in the sample or not? Power law behaviour predicts the existence of large returns and stock market crashes such as the 1929 or 1987 crashes, since they are not outliers.[5]

A power law, with its scale-invariant form, implies that large events are just magnified copies of smaller ones, and arise from the same kinds of causes. [6] For example, this is quite different from the idea, still widely held, that large earthquakes require big triggers. If that were so, then it would be comforting, as you would wait a long time before the next big strike. The statistics and the power law in particular, say otherwise. Both large and small earthquakes occur at random, but importantly with different frequencies.[7] What triggers small and large quakes is precisely the same. From this perspective, it does not make sense to look for special explanations for the largest earthquakes. They are no more special or unusual than the tiny shudders constantly rippling beneath our feet.[8]

Punctuated equilibrium – a 'signature' of complex adaptive systems

The term 'punctuated equilibrium' describes the pattern of alternating calm and storm, or periods of quiescence or stasis interspersed with periods of change.[10] Punctuated equilibrium is contrasted with gradualism, which states that evolution generally occurs by smooth steady and continuous progress.[9] Fossil evidence shows long periods of relatively little change, interspersed with periods of explosive innovation and periods of massive extinction.[10] There are long periods of evolution and incremental adaptation in traditional directions (more of the same); that are interrupted or punctuated occasionally by short bursts of 'quantum' revolutionary realignment, that are sudden, dramatic, and pervasive shifts in direction.[11]

Patterns of punctuated equilibrium show up not just in biological evolution, but in other complex systems ranging from physical avalanches, the crashes of stock markets[10], to changes of political regimes, such as in Egypt with the abrupt and unexpected fall from power

[5] Plerou, Vasiliki; Gopikrishnan, Parameswaran; Gabaix, Xavier; Stanley, H. Eugene (2004). *On the Origin of Power-Law Fluctuations in Stock Prices*. Published in Quantitative Finance 4 (February 2004) C11-C15. arXiv:cond-mat/0403067v1 (cond-mat.dis-nn) 2 Mar 2004.

[6] Buchanan, Mark (2000), op. cit., Pp. 66-88.

[7] Gribbin, John (2004). *Deep Simplicity. Chaos, Complexity and the Emergence of Life.* Allen Lane. An imprint of Penguin Books. London, England. Pp. 138-143.

[8] Buchanan, Mark (2000). *Ubiquity – The Science of history…Or Why the World is Simpler Than We Think.* Weidenfeld and Nicholson. London. England. Pp. 22-42.

[9] Wikipedia – *Punctuated Equilibrium.* http://en.wikipedia.org/wiki/Punctuated_equilibrium Retrieved 07/08/2010.

[10] Beinhocker, Eric D. (2007). *The Origin of Wealth. Evolution, Complexity, and The Radical Remaking of Economics.* Random House Business Books. London, Great Britain. Pp. 173.

[11] Mintzberg, Henry (2007). *Tracking Strategies – Toward a General Theory*. Oxford University Press. Great Britain.   Pp. 202 and 303.

of President Hosni Mubarak after 30 years of rule. Large intermittent bursts have no place in equilibrium systems, but are ubiquitous in history, biology and economics'.[12]

Economic progress shows punctuated equilibrium in action. Schumpeter represents advances in economic reasoning as nonlinear. The world involved much more of discontinuity and seismic convulsions, rather than regularity and smooth progression.[13] Whilst most individual technological innovations have only an incremental impact, overall, technological innovation proceeds in patterns of calm and storm.[14]

'Punctuated equilibrium' describes how science progresses. Thomas Kuhn described the development of science through periods of stable growth punctuated by revisionary revolutions. According to Kuhn the development of science is not uniform but has alternating 'normal' and 'revolutionary' phases. The revolutionary phases are not merely periods of accelerated progress, but differ qualitatively from normal science.[15] Kuhn called the violent revolutionary periods of upheaval 'paradigm shifts'.[16] Discovery of new paradigms in science is the sudden 'punctuation', whilst 'normal' science is the gradual 'equilibrium'.[17] As seen through the lens of the critical state, great revolutions are not necessarily unique in terms of their causes. They are simply the expected large fluctuations of a system poised in a critical state.[18]

Power-law distributions are the fingerprint of self-organized criticality[19]

Self-organized criticality is a property of most non-equilibrium systems, especially where power laws apply.[19] Self-organising criticality automatically drives itself to a critical point by some dynamic process, no matter what state it starts off in. These systems are perpetually out of balance, but organized in a poised state - the critical state – where anything can happen within well-defined statistical laws.[20] Self-organised criticality is the way nature is able to make enormous transformations over short time scales.[21]

[12] Bak, Per (1997). *How Nature Works – The Science of Self-Organized Criticality*. Oxford University Press. Oxford, U.K. Pp. 28-29.

[13] McCraw, Thomas K. (2007). *Prophet of Innovation – Joseph Schumpeter and Creative Destruction.* The Belknap Press of Harvard University Press. Cambridge, Massachusetts, U.S.A. op. cit., Pp. 352, 461-466, 474.

[14] Beinhocker, Eric D. (2007). *The Origin of Wealth. Evolution, Complexity, and The Radical Remaking of Economics.* Random House Business Books. London, Great Britain. Pp. 174-175.

[15] Stanford Encyclopedia of Philosophy. *Thomas Kuhn.* Retrieved 07/08/2010. http:/plato.stanford.edu/entries/Thomas-kuhn/

[16] Beinhocker, Eric D. (2007), op. cit., Pp. Xi.

[17] Kuhn, Thomas S. (1996). *The Structure of Scientific Revolutions*. Third Edition. The University of Chicago Press. Chicago. U.S.A.

[18] Buchanan, Mark (2000), op. cit., Pp. 176-184.

[19] Ball, Philip (2004). *Critical Mass – How one thing leads to another*. Arrow Books. London. United Kingdom. Pp. 283-299.

[20] Ibid., Pp. xi.

[21] Bak, Per (1997), op. cit., Pp. 60-61.

The non-equilibrium master metaphor is the sandpile. Whereas macro-level distributions emerge as stable features of scale-free systems, at the micro-level, such systems exhibit a strong degree of path dependence. To use the sandpile as an illustration, it matters exactly where and when the grains land. This means that point prediction often turns out to be futile, as exemplified by earthquakes. This does not mean, however, that no regularities exist. In particular, it is important to distinguish complex adaptive systems of the self-organised criticality (SOC) kind from mere chaos, which also generates unpredictable behaviour.[22]

What is unique about SOC systems is that they do not have to be carefully tuned to stay at the critical point where they generate the scale-free output responsible for the power laws. They do it automatically, i.e. they are 'self' organised.

The characteristics of non-equilibrium self-organized critical-point phenomena, such as extreme sensitivity to fluctuations, 'scale-free' events and, in particular, power law probability distributions are found widely both in nature and in the world of human affairs. Power-law distributions are the fingerprint of self-organized criticality.[23]

The idea that large systems tend to evolve into a poised 'critical' state, way out of balance, where minor disturbances may lead to events, called avalanches, of any size, can easily be applied to politics, such as looking at events in the Middle East, relations between the United States and China, the oil market, disease, nuclear proliferation, cyber warfare or a dozen other problems of global affairs and security.[24]

Small changes in complex adaptive systems can have a disproportionate effect – the Butterfly effect

Linear thinking assumes that large events need large causes, for example, the large mass extinction that killed the dinosaurs was caused by a very large meteorite. However, non-linear thinking points out those large catastrophic events occur as a consequence of the same dynamics that produces small ordinary events. This observation runs counter to the usual way of thinking about large events, in which we look for equally large causes.[25]

Nonlinearity causes small changes on one level of organisation to produce large effects at the same or different level. This is called positive feedback.[26] The phenomenon where a small change in the initial state of a dynamic system can dramatically alter its subsequent evolution

---

[22] Cederman, Lars-Erik (2002). *Modelling the Size of Wars – From Billiard Balls to Sandpiles.* August 19, 2002. Department of Government, Harvard University. Mass. 02138, U.S.A. Pp. 1-20.

[23] Ball, Philip (2004). *Critical Mass – How one thing leads to another.* Arrow Books. London. United Kingdom. Pp. 283-299.

[24] Ibid., Pp. 49-60.

[25] Bak, Per (1997), op. cit., Pp. 26-32.

[26] Coveney, Peter and Highfield, Roger (1995). *Frontiers of Complexity. The search for Order in a Chaotic World.* Fawcett Columbine. New York. U.S.A. Pp. 7-10.

is called the 'Butterfly effect'.[1] Characterised as the so-called butterfly effect: a butterfly flaps its wings over the Amazon rain forest and sets in motion events that lead to a storm over Chicago. The next time the butterfly flaps its wings, however, nothing of meteorological consequence happens.[27] In non-linear systems small inputs can (but not always) lead to dramatically large consequences. Also, a very slight difference in the initial conditions produces very different outcomes, which is the reason for the unpredictability in weather and other non-linear systems.[27]

Most of the time, small events, small shocks to the system, will only have small impacts, and large shocks will usually have big consequences. But the fact that we observe power-law behaviour in a system tells us that the system operates in a way that means that these relationships do not always hold. Sometimes, a very small event can have profound consequences, and occasionally a big shock can be contained and be of little import. It is not the power law itself which gives rise to these unexpected features of causality; rather it is the fact that we observe a power law in a system which tells us that causality within it behaves in this way.[28]

Once a complex adaptive system is at the critical point and there is tension, small events can have disproportionate consequences. For example, the First World War started from the assassination of the Austro-Hungarian Archduke Franz Ferdinand.[29] It was hardly the first time a member of a ruling family was killed by an assassination and when it does occur this does not usually signal a reason for war. Another example of non-linear change is the unexpected fall of the Soviet Union which demanded a complete remapping of the world. The end of the U.S.S.R. stunned the same minds that were later stunned by 9/11 and the financial collapse of 2008. The U.S.S.R. collapsed quickly.

Scientific knowledge seems to be a SOC system as it tends to be poised in a critical state. If so, then scientists should expect the unexpected. For the fabric of ideas is organised so that the tiniest chance discovery may at any moment and without warning activate a domino-like chain of effects leading to a terrific revolution. Foreseeing such revolutions is next to impossible, for the ultimate consequences of any new idea depends not so much on its own inherent profundity as on where it happens to fall within the network of all scientific ideas. So while chains of events in these systems may not be predictable, it is not the case that nothing is predictable. It is in the statistical pattern that emerges over many chains of events that we can hope to discover the laws for things historical.[30]

---

[27]   Lewin, Roger (1992). *Life at the Edge of Chaos*. Pp. 11.

[28]   Ormerod, Paul (2005). *Why Most Things Fail – Evolution, Extinction and Economics*. Pantheon Books, New York, U.S.A. Pp. vii- 33.

[29]   Ball, Philip (2004), op. cit., Pp. 303-304.

[30]   Buchanan, Mark (2000), op. cit., Pp. 176-184.

Fossil records tell us that extinctions happen on all scales, all the time, and that (like earthquakes) an extinction of any size can happen at any time. Some extinctions might be triggered by meteorite impacts; some might be triggered by Ice Ages. Since power law systems do not need a large trigger in order to set off a large event, an extinction of any size might be set off by a trigger of any size. What matters is that we are dealing with a complex system – life on Earth – which is self-organising, feeding off a flow of energy, and exists at the edge of chaos.[31]

The power law implies that a war, when it starts out, does not 'know how big it will become' – nor does anyone else. The power-law, suggests that the world's political and social fabric tends to be organised on the very edge of instability, and in such a way that wars spread in a special way, so that their ultimate extent is nearly impossible to foresee. The scale free character of the power law suggests that at the outset of a war there are no obvious clues about how large it will become.[32] The stress on SOC encourages us to search for scale-invariant explanations. This means that rather than treating large wars, such as world wars, as qualitatively distinct events that require separate explanations, to get to a unified theory you need to explain all wars regardless of their size.[22]

Systems that have power law distributions have 'emergent' properties or patterns

Emergent properties show that the whole can be more than the sum of the parts[33] and the so-called emergent properties can only be described at higher levels than those of the individual units.[34] The central idea of emergence is the idea that meaningful order can emerge spontaneously, all on its own in complex systems made of numerous interacting parts.[35] An emergent property cannot be deduced or explained in terms of the properties of its constituents. In other words, the whole is greater than the sum of its parts.[1] Feedback loops are a key factor behind this effect.

Examples of emergent properties include:

- The pattern of the **whirlpool** is the result of the dynamic interactions between the individual water molecules. The whirlpool is a collective or '**emergent**' property of the system itself. It is not an expected result, if you just study the properties of a single water molecule sitting in isolation. You need to have a few billion water molecules

---

[31]   Gribbin, John (2004), op. cit., Pp. 151-7.
[32]   Buchanan, Mark (2000). *Ubiquity – The Science of history…Or Why the World is Simpler Than We Think.* Weidenfeld and Nicholson. London. England. Pp. 190-193.
[33]   Ball, Philip (2004), op. cit., Pp. 155.
[34]   Coveney, Peter and Highfield, Roger (1995), op. cit., Pp. 7-10.
[35]   Buchanan, Mark (2002). *Nexus – Small Worlds and the Groundbreaking Science of Networks.* W.W. Norton and Company. New York. U.S.A. Pp. 198.

together, add some energy in the right way, and then you get the complex macro pattern of a whirlpool.[36]

- 'Society' and 'market forces' are emergent phenomena since they are formed from the interactions of millions of people making independent decisions which have different features than those of the studies of individuals that make it up.[37]

- Pareto's law of wealth distribution (power law) is an emergent property.[35] Producing a power law is an emergent result as it is by no means the only relationship that can happen, e.g. it could have been a bell curve. There is nothing pre-ordained in producing a power law. Everyone is completely free to decide their own actions, yet out of this multiplicity of individual choices comes a mathematical law that holds true.[38]

- Life is an emergent property, as are human beings.[39] Zipf's law of cities or words is an emergent property in the sense that it is not an obvious consequence of the underlying dynamical rules.[40]

Power law distributions means the 'the winner takes all' Matthew effect applies

The Pareto process is a rich-get-richer mechanism in which the most populous cities or best-selling books get more inhabitants or sales in proportion to the numbers they already have. These distributions have a power law in their tail. These 'rich-get-richer' processes are also known as the Matthew effect, cumulative advantage or preferential attachment. The 'Matthew' effect comes from the biblical edict, 'For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath.' (Matthew 25:29 – King James Bible). This edict is harsh, but is very apt for reality. The Matthew effect helps explain why only a very few have most of the citations, only a few have most of the wealth, why a few cities have most of the population, why a few sites on the World Wide Web are so dominant, why only a few entertainers and film stars earn so much more that others, why only a few are so sexually active by number of partners, as well as why only a few are so well connected in social networks. When studying power law distributions, studying the key few is important, since the extreme power law outliers, sometimes called 'aristocrats', are so well connected (or wealthy etc) that their activity dominates all else and no unfocussed attention has any chance of success.[41]

---

[36] Beinhocker, Eric D. (2007), op. cit., Pp.18.
[37] Beinhocker, Eric D. (2007), op. cit., Pp. 450.

[38] Ball, Philip (2004), op. cit., Pp. 479-501.
[39] Coveney, Peter and Highfield, Roger (1995), op. cit., Pp. 330.
[40] Bak, Per (1997), op. cit., Pp. 26-32.
[41] Buchanan, Mark (2002), op. cit., P182-183.

Book sales are an example of the Matthew effect, since book sales have the typical long tail power law distribution. For example, only 10 titles (out of 1.2 million) in 2004 sold more than a million copies, whereas 1.15 million titles sold less than a 1,000 books each. In other words, about 98 percent of books are uncommercial whether they were intended that way or not.

Pareto's Law of Wealth provides the following examples of this uneven wealth distribution:

- The world's richest 2 percent own more than 50 percent of the world's wealth, whilst the poorest 50 percent of people own 1 percent.[42]

- In the United States only 20 percent of the population of the country own 85 per cent of the wealth.[3] In the U.S.A. the benefits of economic growth over the last 30 years have largely been concentrated in the top 1%.[43]

- In New Zealand the top 10% of the population holds 50% of the total wealth, and the bottom half owns around 5%. This compares with other countries like the US, UK, Australia and Canada where the bottom 30% of the population have less than 1% of the wealth.[44]

This contrasts with what would happen if a normal Gaussian distribution applied to cities, book sales or wealth, where there would be a tight clustering around a mean. This makes an 'unreal' world of essentially one size of city, one size of book sales and a very even wealth distribution. This is 'unreal' as it has never occurred at any time or location in history. For these non-linear phenomena (wealth, book sales etc) it is inherently an unfair non-linear world.

Change is triggered by both internal/ external reasons

Traditional economics and evolution believed that different extinctions are caused solely by shocks that are external to the system. However, for complex adaptive systems such as species extinctions and for the major ups and downs in the sharemarket, there are not always obvious external shocks than match dramatic changes in those systems.[45]

Complexity theory believes that extinctions of all sizes (for species or companies) and sharemarket movements arise quite naturally from within the system itself, from the network of connections between the individual agents and how they evolve over time.[45] The idea of needing an 'external' force is essentially a 'linear' idea, since it assumes that a large effect must come from a large impact. In a non-linear world, there is no simple way to predict

---

[42] Ball, Philip (2004), op. cit., Pp. 307-8.
[43] Wikipedia – *Pareto principle*. http://en.wikipedia.org/wiki/80/20_rule 04/05/2009
[44] Bennett, Neville (2009). *State of the Nation's Wealth.* National Business Review. Pp. 10-12 Wealth Guide.
[45] Ormerod, Paul (2005), op. cit., Pp. 169-210.

emergent behaviour. Self-organising criticality does not need a large change to make a large impact, as it is not a linear effect.[46] This is a difficult idea to accept, since society still searches for big causes of share crashes etc.

In areas where complex adaptive systems apply (that is any system where a power law distribution applies) it is now generally believed that both internal/ endogenous triggers and external/ exogenous triggers can be in operation at the same time. The two approaches to causing triggers are complementary and do not exclude each other.[47]

Power law distributions give pattern prediction not point prediction

We tend to make predictions by looking for patterns in the past and then extrapolating those patterns out into the future. Often this works quite well, especially if the world is fairly stable. But the frequent occurrence of 'stable' periods between punctuations points only gives us a false sense of predictability since the existence of power laws means that big changes are far more likely to occur than the expectations our mental models give us on the basis of past experience.[48]

Friedrich Hayek's main contribution to early complexity theory was to point out that you could only explain complex phenomena from pattern prediction, compared with the precise predictions in simple linear systems like classical physics.[49]

Self-organizing criticality only tells you about the overall statistics of avalanches; it tells you nothing about any particular avalanche. This is another case where understanding is not the same thing as prediction.[50] For example, knowing the history of events in non-linear systems often provides little or no prediction value as to the next event, despite the fact that each step follows logically from the previous step (historical explanations looking backwards).[51]

For example, you will never be able to predict the size of the next earthquake or war, but the power law graph for that subject gives a clear forecasting framework at the highest level looking at all the available data, which we can learn from and help us get over the idea of needing to predict the next event.

---

[46]   Bak, Per (1997), op. cit., Pp. 130-131.

[47]   Gould, Stephen Jay (2007). *Punctuated Equilibrium.* The Belknap Press of Harvard University Press. Cambridge, Massachusetts. U.S.A. Pp. 235.

[48]   Beinhocker, Eric D. (2007), op. cit., Pp. 328.

[49]   Wikipedia – *Complex Systems.* http://en.wikipedia.org/wiki/Complex_systems#Complexity_and_Chaos_theory 07/05/2010

[50]   Waldrop, M. Mitchell (1992). *Complexity – The Emerging Science of Order and Chaos.* Simon and Schuster. New York. U.S.A. Pp. 306.

[51]   Bak, Per (1997), op. cit., Pp. 60-61.

Another example is Pareto's law of wealth which is not about individuals. Rather, it captures a pattern that emerges at the level of many individuals and leaves individual histories aside. Pareto's law reflects little about human cultural, behavioural, or intellectual features and arises instead as a consequence of some deeper principle of organisation.[52]

The signature of a power law distribution is the straight line 'log-log' graph

All power laws tend to look alike when plotted on a graph with a logarithmic scale. (Where each division is a factor of 10 larger than the one that came before it, for example, 1, 10, 100, 1,000 and so on.) Common examples of logarithmic scales are the Richter scale of earthquakes and the decibel scale of sound volume. When a power law is plotted on a log scale on both axes (called a 'log-log graph'), a straight line sloping down should result. The exact angle of that slope varies, but whether it's the size and frequency of earthquakes or the distribution of publicly listed companies by stock-market capitalization, the slope is always a straight line.[53]

The exponent of the slope is called the 'power' which is why we use the term power law. Each power law is characterized by a unique exponent. The power law describes the degree distribution and the exponent is often called the degree exponent. For example, the distribution of incoming links on web pages follows a power law with a degree exponent close to two. The collective action of millions of webpage creators work together in a way that produces this emergent result, that is not a 'random' (i.e. not a normal distribution) network.[54] The exponent is therefore an important and powerful forecasting tool.[55]

The straight line log-log graph is often called the signature of a power law. However, whilst this is a necessity, there are non-power data sets (such as 'log-normal' or exponential) that mimic this signature behaviour, so care needs to be taken.[56]

High level power-law patterns (from straight line log-log power law charts)

Examples include:

- *Authors (Lotka's law)* - As the number of articles published increases, authors producing that many publications become less frequent. For example, there are ¼ as many authors publishing two articles within a specified time period as there are

---

[52] Buchanan, Mark (2002), op. cit., P189 -190.
[53] Anderson, Chris (2006). *The Long Tail – Why the Future of Business is Selling Less of More.* Hyperion. New York. U.S.A. Pp. 126-7.
[54] Barabasi, Albert-Laszlo (2002), op. cit., Pp. 67-72.
[55] Taleb, Nassim Nicholas (2007). *The Black Swan – The Impact of the Highly Improbable*. Penguin Books. London. England. Pp. 262-5.
[56] Wikipedia – Power Law. http://en.wikipedia.org/wiki/Power_law 07/05/2010

single-published authors, 1/9 as many publishing three articles, 1/16 as many publishing four articles, etc.[57]

- *Company sales* - This means that firms with sales of $1 million are four times as numerous as those with $2 million, which are in turn four times as numerous as those with $4 million.[58] Scaling means firms of size 1,000 are four times less common than firms of size 500, while firms of size 10,000 are 100 times less common than firms of size 1,000.

- *Earthquake energy (The Gutenberg/ Richter scale)* - An increase of one unit on the scale corresponds to an increase in the amount of energy released by a factor of 30; a magnitude 2 earthquake is 30 times as powerful as a magnitude 1 earthquake, a magnitude 3 earthquake is 30 times more powerful than a magnitude 2 earthquake (and therefore 900 times more powerful than a magnitude 1 earthquake), and so on.[59] In terms of energy, the data indicates that whenever you double the energy released by an earthquake, it is one quarter less frequent.[60]

- *Earthquake magnitude (The Gutenberg/ Richter scal*e) - For every 1,000 earthquakes of magnitude 5 there are roughly 100 earthquakes of magnitude 6, 10 earthquakes of magnitude 7, and so on.[59]

- *Evolution* - The fossil records show the size of extinction events, with a power law exponent close to 2. The distribution of the lifetimes of species also appears to follow a power law, with an exponent of about 1.7. The number of species per genus also appears to follow a power law with an exponent of about 1.5.

- *River networks* - Every time you double the amount of water drained, you find that the number of such streams falls by a factor of about 2.7. If there are one hundred streams draining a thousand square kilometres, then there will be only about thirty-seven draining two thousand square kilometres, and so on.[61]

- *Scientific citations (Redner's law)* - Double the number of citations and the number of papers receiving that many falls off by about eight.[62]

---

[57] Wikipedia – *Lotka's Law*. http://en.wikipedia.org/wiki/Lotka%27s_law 07/05/2010
[58] Buchanan, Mark (2007). *The Social Atom – Why the Rich get Richer, Cheaters Get Caught, and Your Neighbour Usually Looks Like You*. Bloomsbury. New York. U.S.A. Pp. 180-186.
[59] Gribbin, John (2004), op. cit., Pp. 138-143.
[60] Buchanan, Mark (2000), op. cit., Pp. 22-42.
[61] Buchanan, Mark (2007), op. cit., Pp. 174-177 and 228.
[62] Buchanan, Mark (2000), op. cit., Pp. 180-184.

- *Share market returns and numbers of trades -* The power law exponent is 3 (cubic law of returns). The cubic law predicts the existence of a few very large returns over a century.[5]

- *Size of cities (Zipf's law)* – For example a rank 3 city would have 1/3 the population of a country's largest city, a rank four city would have ¼ the population of the largest city, and so on.[63]

- *The World Wide Web* **-** the number of nodes having a certain number of links decreased by about five each time the number of links was doubled. In these networks, just a few nodes have so many links that 80 to 90 percent of the network's total number of links feed into just a small fraction of the nodes - an architecture that is dominated by especially well connected hubs.[64]

- *War (Richardson's law)* - Every time you double the number of deaths, the wars of that size become four times less common. If you use the number of casualties divided by the world population at the time, the precise pattern becomes 2.62 times less frequent every time you double the number of deaths.[65]

- W*ealth (Pareto's law)* - Each time wealth goes up by a factor of ten, the number of people having that much falls by a factor of about six (about six times as many people have $1 million, for example, as have $10 million).[66] Said another way, if wealth doubles from $1 million to $2 million, the incidence of people with at least that much money is cut in four, which is an exponent of two.[67]

For the emergent power law pattern, the details don't matter

Zipf's power law distribution of company sizes comes out of all the chaos of millions of people doing an infinite variety of different things, from which emerges a mathematical pattern.[68] The skewness appears to be insensitive to changes in political and regulatory environments, immune to waves of mergers and acquisitions, and unaffected by surges of new firm entry and bankruptcies. Whilst the overall emergent pattern remains constant over time, at the individual level of firms there is perpetual churn. The other key point to note is the fact that details don't matter.[69] Prediction of specific events and details is impossible.[68]

---

[63]   Wikipedia, the free encyclopedia (2009). *Rank-size distribution*. http://en.wikipedia.org/wiki/Rank-size_Law
[64]   Buchanan, Mark (2002), op. cit., Pp. 20 and 83-85.
[65]   Buchanan, Mark (2000), op. cit., Pp. 190-193.
[66]   Buchanan, Mark (2007), op. cit., Pp. 168-9.
[67]   Taleb, Nassim Nicholas (2007), op. cit., Pp. 262-5.

[68]   Buchanan, Mark (2007), op. cit., Pp. 180-186.
[69]   Axtell, Robert L. (2001). *Zipf Distribution of U.S. Firm Sizes*. Volume 293 Science, 7 September 2001. www.sciencemag.org Pp. 1818 – 1820.

Power law patterns offer an aid in getting to grips with systems that simply never seem to settle down, that are always churning and changing and evolving, where nothing seems constant, and nothing is ever repeated.[68]

Increased networks/ interdependence, means increased turbulence

Networks are an essential ingredient in any complex adaptive system.[70] Unfortunately this means that many of the most important world networks – economic, political systems, ecosystems, and so on – are poised perpetually on the very edge of instability and tumultuous upheaval. As a consequence, it is something akin to a universal law of nature that the course of history must necessarily be punctuated – and quite frequently – by seemingly inexplicable upheavals. The character of history will therefore be long periods of relative calm and gradual change to be punctuated by staggeringly overwhelming events that totally reshape the social and political landscape.[71]

There exists a dangerous paradox – the more closely we're bound, the less resilient we all become. Studies of food webs or trade networks, electrical systems and stock markets, find that as they become more densely linked they also become less resilient; networks, afterall, propagate and even amplify disturbances. Worse, the more efficient these networks are, the faster they spread those dangers. Drop a shock into a network and you get a chain reaction and the end state can not be predicted.[72] The turbulence in all these systems dramatically increases due to power law effects within network flows. Extremes will get more extreme due to the 'Matthew' effect, where the winners take all. All these interlocking economic/ technology networks or webs are subject to cascades of change leading to the emergent pattern of punctuated equilibrium.[73]

Strategy is about surviving the 'Red Queen Race'

Schumpeter said that, 'Strategy must be seen in its role in the perennial gale of creative destruction; it can-not be understood irrespective of it or, in fact on the hypothesis that there is a perennial lull'. Schumpeter believed that a creative response as opposed to an adaptive response can never be predicted and is therefore indeterminate, shaping long-run outcomes in a country, industry, or firm.[74] The punctuated nature of change tricks our pattern-recognizing minds into thinking that the world is more stable than it really is.[75]

[70] Beinhocker, Eric D. (2007), op. cit., Pp. 141.
[71] Buchanan, Mark (2002), op. cit., Pp. 20 and 83-85.
[72] Ramo, Joshua Cooper (2009). The Age of the unthinkable – Why the new world disorder constantly surprises us and what we can do about it. Little, Brown. London U.K. Pp. 49-60.
[73] Beinhocker, Eric D. (2007), op. cit., Pp. 174-175.
[74] McCraw, Thomas K. (2007), op. cit., Pp. 352, 461-466, 474.
[75] Beinhocker, Eric D. (2007), op. cit., Pp. 328-334.

There is no such thing as a safe, stable industry in the long term. Schumpeter's ghost is alive and well and the 'gales of creative destruction' are blowing harder than ever. Competitive advantage is rare and short-lived in the biological world as well. In biological systems, species are locked in a never-ending co-evolutionary arms race (prey v predator) with each other. Biologists refer to such co-evolutionary spirals as Red Queen races, named after the Red Queen in Lewis Carroll's Through the Looking Glass. It was she who said, 'In this place it takes all the running you can do, to keep in the same place'. There is no such thing as winning a Red Queen race, the best you can ever do is run faster than the competition. In evolutionary systems, sustainable competitive advantage does not exist; there is only a never-ending race to create new sources of temporary advantage. Understanding this means we change our definition of an excellent company from one of continuous high performance for very long periods (an achievement that is almost non-existent) to one that can string together a series of temporary advantages over time – in other words, a strong runner in the Red Queen race.[75] Learning faster than your competitors is the only sustainable advantage in an environment of rapid innovation and change. Resilience is key to survival.

Importance of modelling

One very important tool which I have hardly mentioned in this thesis is 'modelling'. Probably every power-law example in this thesis comes from sources that not only record the power-law log-log chart but also record attempts at modelling that subject using that chart's power-law co-efficient. I acknowledge the extreme importance of power-law charts to modellers since this has provided new methods and a greatly enlarged range of subjects (i.e. non-linear) to model. As stated in the introduction, my interest is only in the high level principles of non-linearity as it affects strategy. I am not a modeller, so I have not emphasized that aspect but that should not be taken as a lack of appreciation of how important power-laws are to modelling over a whole range of subjects. Whilst 'modelling' is an important technique it is not the one word answer to my thesis question. It is not my interest or my journey for this thesis. However, I believe this thesis would be a good background explanation of the subject of power-law non-linearity for modellers and non-modellers (like me).   It doesn't help you 'predict' the next event, but it should provide a good description of what is happening or could happen.

For every power-law subject contained in this thesis, modellers are working on 'bottom-up' modelling where they focus on the lower-level individual entities or agents that make up the system. From my literature review, I can conclude that power–laws with their straight line graphs (and therefore power-law coefficients) have provided a wonderful tool to modellers to increase their understanding of the world (and therefore ours) and to understand the dynamics of agents as never before. Non-linear systems were previously considered too difficult to understand and predict. Now, with the techniques described in this thesis, the study of non-linear systems is becoming more productive.

Now that I have completed this study of power laws and non-linearity, I believe I have found a useful framework (Complexity) to understand strategy. For example, when dealing with linear matters, linear based strategic planning approaches are appropriate, but when dealing with non-linear matters, then non-linear strategic planning approaches are essential. Understanding non-linearity is important in understanding such things as 'Black Swans' and 'punctuated equilibrium. Luck, serendipity and 'bounded rationality' are always involved in complex adaptive systems, whereas linear systems tend to comply with the so called 'rational' traditions in science and economics. The strategist needs a strategy toolkit that can handle both linear and non-linear situations.

Final thoughts – No difference between natural and human power law activities

For this thesis I have found no effective difference between natural effects (e.g. earthquakes or animal extinction rates) that follow power law distributions and human activities (e.g. wealth, war, citations etc) that follow power law distributions. I have therefore not separated the two different types in this thesis. In fact the later section on networks is in purely alphabetical order of headings. The implication of this lack of clear difference between natural and human effects is both a surprise and is troubling. For example, empirically, the power law for war is a very good fit and that is not a comforting finding. It poses the profound question, as to whether we can actually change this underlying pattern or are we predestined to fit the various power law patterns such as in war or wealth. Are human agents inherently just grains of sand in the 'sandpile' of life or can the grains 'think' and determine their own future? My thesis therefore suggests that 'overall', humans are living in a completely deterministic world, hence power law effects happen at the macro level. Yet as individuals at a micro level, we like to think that human agency does count. Life and strategy is therefore all about striving to do better in a world that is often very unfair and non-linear. Understanding the rules of life is therefore essential.

Bibliography

Anderson, Chris (2006). *The Long Tail – Why the Future of Business is Selling Less of More.* Hyperion. New York. U.S.A.

Anderson, Philip W. (1997) *Some Thoughts about Distribution in Economics* – Part of *The Economy as an Evolving Complex System II.* (Editors Arthur, W. Brian; Durlauf, Steven N.; Lane, David A) - Proceedings Volume XXVII. Santa Fe Institute. Studies in the Sciences of Complexity. Addison-Wesley. The Advanced Book Program. Reading Massachusetts. Pp. 565-6

Attali, Jacques (2009) *A Brief History of the Future. A brave and controversial look at the twenty-first century.* Allen and Unwin, Crowns Nest, NSW.

Australian Government/ Australian Public Service Commission (2007). *Tackling Wicked Problems – A Policy Perspective.* Contemporary Government Challenges. http://apsc.gov.au/publications07/wickedproblems.pdf

Axtell, Robert and Florida, Richard (March 2006). *Emergent Cities: Micro-Foundations of Zipf's Law.* Brooking Institution. U.S.A.

Axtell, Robert L. (2001). *Zipf Distribution of U.S. Firm Sizes*. Volume 293 Science, 7 September 2001. www.sciencemag.org Pp. 1818 – 1820.

Axtell, Robert (June 1999). *The Emergence of Firms in a Population of Agents: Local Increasing Returns, Unstable Nash Equilibria, and Power Law Size Distributions.* Centre on Social and Economic Dynamics, Working Paper No. 3. June 1999. Brookings Institution. U.S.A.

Bak, Per (1997). *How Nature Works – The Science of Self-Organised Criticality.* Oxford University Press. Oxford. U.K.

Ball, Philip (2004). *Critical Mass – How one thing leads to another.* Arrow Books. London. United Kingdom.

Barabasi, Albert-Laszlo (2002). *Linked – The New Science of Networks.* Perseus Publishing. Cambridge, Massachusetts, USA.

Beinhocker, Eric D. (2007). *The Origin of Wealth. Evolution, Complexity, and The Radical Remaking of Economics.* Random House Business Books. London, Great Britain.

Bennett, Neville (2009). *State of the Nation's Wealth.* National Business Review. Wealth Guide.

Buchanan, Mark (2002). *Nexus – Small Worlds and Groundbreaking Science of Networks.* W.W. Norton & Company. New York. U.S.A.

Buchanan, Mark (2007). *The Social Atom – Why the Rich get Richer, Cheaters Get Caught, and Your Neighbour Usually Looks Like You.* Bloomsbury. New York. U.S.A.

Buchanan, Mark (2000). *Ubiquity – The Science of history…Or Why the World is Simpler Than We Think.* Weidenfeld and Nicholson. London. England.

Casti, John L (1997). *Would-Be Worlds. How Simulation is Changing the frontiers of Science.* John Wiley and Sons, Inc. New York. U.S.A.

Cederman, Lars-Erik (2002). *Modelling the Size of Wars – From Billiard Balls to Sandpiles.* August 19, 2002. Department of Government, Harvard University. Mass. 02138, U.S.A.

Clauset, Aaron, Shalizi, Cosma Rohilla, and Newman M.E.J. *Power-Law Distributions in Empirical Data.* arXiv:0706.1062v2 (physics.data-an) 2 Feb 2009 www-personal.umich.edu/-mejn/pubs.html (Publications of Mark Newman) University of Michigan.

Cohen, Jack and Stewart, Ian (1994). *The Collapse of Chaos. Discovering Simplicity in a Complex World. Viking.* Published by Penguin Group. Great Britain.

Coveney, Peter and Highfield, Roger (1995). *Frontiers of Complexity. The search for Order in a Chaotic World.* Fawcett Columbine. New York. U.S.A.

Dator, James A. (2002). *Advancing futures - Futures Studies in Higher Education*. Chapter 16, *Teaching Futures-Seeking Policy process* by Markku Sotarauta. Praeger Studies on the 21st Century. Westport, Connecticut. United States of America.

Foster, Richard and Kaplan, Sarah. (2001) *Creative Destruction – Why Companies that are built to last underperform the Market – and how to successfully transform them*. A Currency Book. Doubleday. New York. U.S.A.

Foxman, Betsy; Newman, Mark; Percha, Bethany; Holmes, King K; Aral, Sevgi O. (2004?) *Measures of Sexual Partnerships: lengths, gaps, overlaps and Sexually Transmitted Infection*. Pp1-23. University of Michigan. U.S.A.

Friedman, George (2009). *The Next 100 Years – A Forecast for the 21st Century*. Doubleday, New York.

Friedman, Thomas L. (2008). *Hot, Flat, and Crowded – Why we need a green revolution and how it can renew America*. Allen Lane, London.

Friedman, Thomas L. (2005). *The World is Flat – A Brief history of the twenty-first century.* Farrar, Straus and Giroux, New York.

Gabaix, Xavier; Gopikrishnan, Parameswaran; Plerou, Vasiliki; Stanley, H. Eugene (2004). A *Theory of Large Fluctuations in Stock Market Activity*. February 17, 2004.

Gilbert, Natasha (2009). *Modellers claim wars are predictable – Insurgent attacks follow a universal pattern of timing and casualties*. December 16, 2009. www.scientificamerica.com/article.cfm?id=modellers-claim-wars-are-predi&print=true

Glenn, Jerome C. and Gordon, Theodore J. and Florescu, Elizabeth. *2007 State of the Future*. Millennium-project.org Accessed 1 September 2009.

Gribbin, John (2004). *Deep Simplicity. Chaos, Complexity and the Emergence of Life.* Allen Lane. An imprint of Penguin Books. London, England.

Gould, Stephen Jay (2007). *Punctuated Equilibrium*. The Belknap Press of Harvard University Press. Cambridge. Massachusetts. U.S.A.

Gould, Stephen Jay (1989). *Wonderful Life. The Burgess Shale and the Nature of History*. Hutchinson Radius. London. U.K.

Hamel, Gary (2000). *Leading the Revolution*. Ch.9: The New Innovation Solution. Harvard Business School Press, Boston.

Hamel, Gary and Prahalad, C.K. (1996). *Competing for the Future*. Harvard Business School Press. Boston, Massachusetts.

Johnson, Neil F.; Spagat, Mike; Restrepo, Jorge A.; Becerra, Oscar; Bohorquez, Juan Camilo; Suarez, Nicolas; Restrepo, Elvira Maria; and Zarama, Roberto. (2005) *Universal patterns underlying ongoing wars and terrorism.* http://xxx.lanl.gov/abs/physics/0506213 Oxford, London and Bogota, Colombia.

Kauffman, Stuart A (1995). *At Home In The Universe – The Search for the Laws of Self-Organization and Complexity.* Oxford University Press. New York. U.S.A.

Kauffman, Stuart A (2000). *Investigations.* Oxford University Press. New York. U.S.A.

Krugman, Paul (1997). *How the Economy Organizes itself in Space: A Survey of the New Economic Geography* – Part of *The Economy as an Evolving Complex System II.* (Editors Arthur, W. Brian; Durlauf, Steven N.; Lane, David A) - Proceedings Volume XXVII. Santa Fe Institute. Studies in the Sciences of Complexity. Addison-Wesley. The Advanced Book Program. Reading Massachusetts.

Kuhn, Thomas S. (1996). *The Structure of Scientific Revolutions*. Third Edition. The University of Chicago Press. Chicago. U.S.A.

Lawrence, Eton (1999). *Strategic Thinking: A Discussion Paper.* Research Directorate. Policy, Research and Communications Branch. Public Service Commission of Canada.

Levy, Jack S. (1983). *War in the Modern Great Power System 1495-1975.* The University Press of Kentucky. Lexington, Kentucky, U.S.A.

Levy, Jack S. (1982). *The Contagion of Great Power War Behaviour, 1495-1975.* American Journal of Political Science, Vol.26, No. 3. (Aug., 1982), pp. 562-584. University of Texas Press. U.S.A.

Levy, Jack S. and Morgan, T. Clifton (1984). *The Frequency and Seriousness of War: An Inverse Relationship?* The Journal of Conflict Resolution, Vol. 28, No. 4. (Dec., 1984), pp. 731-749. Sage Publications, Inc. U.S.A.

Lewin, Roger (1992). *Life at the Edge of Chaos*.

Mandelbrot, Benoit B. and Hudson, Richard L. (2008). *The (mis)Behaviour of Markets – A Fractal View of Risk, Ruin, and Reward.* Profile Books. London. Great Britain.

Martin, Roger L. (2007) *The Opposable Mind – How Successful Leaders Win Through Integrative Thinking.* Harvard Business School Press. Boston, Massachusetts.

Matia, Kaushik; Ashkenazy, Yosef; and Stanley, H. Eugene (2003). *Multifractal Properties of Price Fluctuations of Stocks and Commodities.* arXiv:cond-mat/0308012v1 (cond-mat.stat-mech) 1 Aug 2003.

McCraw, Thomas K. (2007). *Prophet of Innovation – Joseph Schumpeter and Creative Destruction.* The Belknap Press of Harvard University Press. Cambridge, Massachusetts, U.S.A.

Mintzberg, Henry (2007). *Tracking Strategies – Toward a General Theory.* Oxford University Press. Great Britain.

Mintzberg, Henry and Ahlstrand, Bruce and Lampel, Joseph. (2005). *Strategy Bites Back- It is a lot more, and less, than you ever imagined…*Prentice Hall. Pearson Education Limited. Harlow, England.

Mintzberg, Henry (1996) *Crafting Strategy*. Reading 5.2 in the *Strategy Process: Concepts, Contexts, Cases.* Prentice Hall, New Jersey.

Mintzberg, Henry (1994). *The Rise and Fall of Strategic Planning*. Ch 5: Fundamental Fallacies of Strategic Planning. The Free press. McMillan. New York.

Mlodinow, Leonard. (2009).*The Drunkard's Walk*. *How Randomness rules our Lives*. Penguin Books. London. UK.

Newman, M.E.J (2006). *Power laws, Pareto distributions and Zipf's law*. arXiv:cond-mat/0412004v3 (cond-mat.stat-mech) 29 May 2006.

Newman, M.E.J., Clauset, Aaron and Salizi, Rohilla (2009). *Power-Law Distributions in Empirical Data*. arXiv:0706.1062v2 (physics.data-an) 2 Feb 2009.

Newman, M.E.J. (2003). *The structure and function of complex networks.* arXiv:cond-mat/0303516v1 (cond-mat.stat-mech) 25 March 2003.

Newman, Mark; Barabasi, Albert-Laszlo; Watts, Duncan J. (2006). *The Structure and Dynamics of Networks.* Princeton University Press. Princeton. New Jersey. U.S.A.

Newman, M.E.J and Palmer, R.G. (2003). *Modelling Extinction.* A Volume in the Santa Fe Institute Studies in the Sciences of Complexity. Oxford University Press. New York. U.S.A.

Newman, M.E.J (2008). *The First-Mover advantage in Scientific Publications*. arXiv:0809.0522vl (physics.soc-ph) 2 Sep 2008.

Ormerod, Paul (2005). *Why Most Things Fail – Evolution, Extinction and Economics.* Pantheon Books, New York, U.S.A.

Plerou, Vasiliki; Gopikrishnan, Parameswaran; Gabaix, Xavier; Stanley, H. Eugene (2004). *On the Origin of Power-Law Fluctuations in Stock Prices*. Published in Quantitative Finance 4 (February 2004) C11-C15. arXiv:cond-mat/0403067v1 (cond-mat.dis-nn) 2 Mar 2004.

Ramo, Joshua Cooper (2009). *The Age of the unthinkable – Why the new world disorder constantly surprises us and what we can do about it.* Little, Brown. London U.K.

Rosenberg, Nathan (2004). *Innovation and Economic Growth.* O.E.C.D.

Schumpeter, Joseph A. (2009) *Can Capitalism Survive? Creative Destruction and the Future of the Global Economy.* First published 1942. First Harper Perennial Modern Thought. New York. U.S.A.

Small, John and Sweetman, Chris (September 2008). *Drivers of Economic Growth in Auckland. A Report for the Royal Commission on Auckland Governance*. Covec Limited.

Stanley, H.E.; Amaral, L.A.N.; Gopikrishnan, P.; Plerou, V. (2000). *Scale invariance and universality of economic fluctuations. Boston University.* Physica A 283 (2000) Pp. 31-41. Elsevier Science B.V.

Stanford Encyclopedia of Philosophy. *Thomas Kuhn.* Retrieved 07/08/2010. http:/plato.stanford.edu/entries/Thomas-kuhn/

Sterman, John (2002). *All Models are Wrong: Reflections on Becoming a Systems Scientist.* In Systems Dynamics Review. Vol. 18, No. 4 (Winter 2002).

Taleb, Nassim Nicholas (2007). *The Black Swan. The Impact of the Highly Improbable.* Random House. New York.

Taleb, Nassim Nicholas (2007). *Fooled by Randomness. The Hidden Role of Chance in the Markets.* Penguin Books. London. England.

Thuan, Trinh Xuan (2001). *Chaos and Harmony. Perspectives on Scientific Revolutions of the Twentieth Century.* Oxford University Press. New York. USA. Pp.

Waldrop, M. Mitchell (1992). *Complexity – The Emerging Science of Order and Chaos.* Simon and Schuster. New York. U.S.A.

Wallerstein, Immanuel (2003). *The Decline of American Power – The U.S. in a chaotic World.* New Press, New York.

Wikipedia – *Aftershock*. http://en.wikipedia.org/wiki/Aftershock#Omori.27s_Law

Wikipedia – *Complex Systems*. http://en.wikipedia.org/wiki/Complex_systems#Complexity_and_Chaos_theory

Wikipedia – *Lotka's Law*. http://en.wikipedia.org/wiki/Lotka%27s_law

Wikipedia - *Nonlinear system.* http://en.wikipedia.org/wiki/Nonlinear_system

Wikipedia - *Pareto distribution*. http://en.wikipedia.org/wiki/Pareto_distribution

Wikipedia – *Pareto principle*. http://en.wikipedia.org/wiki/80/20_rule

Wikipedia – *Power Law*. http://en.wikipedia.org/wiki/Power_law

Wikipedia – *Punctuated Equilibrium.* http://en.wikipedia.org/wiki/Punctuated_equilibrium

Wikipedia - *Rank-size distribution*. http://en.wikipedia.org/wiki/Rank-size_Law

Wikipedia – Self Organised criticality. http://en.wikipedia.org/wiki/Self-organised_criticality

Wikipedia - *Zipf's law*. http://en.wikipedia.org/wiki/Zipf%27s_law

Zakaria, Fareed (2008). *The Post American World.* W.W. Norton, New York.