

## The Quantification of Systemic Risk and Stability: New Methods and Measures

Romney B. Duffey, Ph.D, B.Sc., FASME

### Abstract

We address the question of the prediction of large failures, or system collapse, and the necessary concepts related to risk quantification, minimization and management. This requires a new approach as predictions using standard statistical techniques and distributions fail in the presence of learning effects, and we have the possibility of rare or low frequency events due to human involvement.

Failures are due to the improper and incomplete treatment of human error as part of the overall system. All major events and disasters, including financial ones, include the dominant contribution not only from individual mistakes, but also management failures and corporate-wide and regulatory errors and blunders. Risk is a measure of our uncertainty, and that uncertainty is determined by the probability of error. We must also estimate and predict risk that also includes the unknown or rare event. Here we unify financial risk concepts with the outcomes, trends and measures already derived for other technical systems.

Obviously, as humans, we learn from experience, both good and bad. Correct measure(s) for experience enable the prediction and uncertainty estimation for the entire range of rare, repeat and unknown outcomes (e.g., major industrial disasters, facility accidents and explosions, every day auto accidents, aircraft crashes and market collapses). We also take risks and must make mistakes in order to improve. A universal curve is derived for both collective and individual learning trends, naturally including the inevitability of outcomes and risk. Based on our work studying and analyzing over 200 years of real data on and for risk in technological, medical, industrial and financial systems, five measures are presented and discussed for the objective measure of risk, failure probability and risk exposure. We also introduce and present the unifying concept of risk and uncertainty derived from the information entropy as a quantitative measure of randomness and disorder. We show how this allows comparative risk estimation and the discerning of insufficient learning. Since these risk measures and learning trends have been largely derived from data including the financial arena, we show how to generalize these to include the presence of market pressures, financial issues and risk measures. We define and present the bases, analyses and results for new risk measures for the quantitative *predictions* of risk exposure, failure and collapse using relevant experience including:

- a) Universal Learning (ULC), similar to the Black-Scholes concept;
- b) Risk Ratios (RR) and exposure, as derived from empirical hazard curves;
- c) Repeat Event Predictions (REP) or 'never happening again', equivalent to birthday matching and re-occurring echoes;
- d) Rare and Unknown Outcome occurrences (UU), as in the black swan concept;
- e) System and Organizational Stability (SOS) or resilience criteria, using the information entropy concept.

We provide quantified examples for production processes, transportation losses, major hazards and financial exposure. These new concepts also provide the probability of success, the emergence of order and the understanding and quantification of risk perception.

### **Risk: How we learn from experience and what we know about risk prediction**

What is the risk of system failure? To answer that question we must understand how and why systems fail, and how to make a prediction.

We first review what is known and not known about predicting and managing risk in industrial, energy, transportation, nuclear, medial, manufacturing systems, and the risk measures. We address the question of the predictability of a large systems failure, or collapse, and the necessary concepts related to risk quantification and system stability that are emerging from the physical sciences, cognitive psychology, information theory and multiple industrial arenas that are relevant to current financial and economic market and stability concerns. We have defined the risk of any outcome (being a proposition of which one is uncertain) as caused by uncertainty, and that the measure of the uncertainty is probability,  $p$ . We attempt to use some of these risk concepts, learning and applications from mainly operational systems to inform financial systems.

Both large and small risks are due to the probability/possibility of an adverse event, outcome, or accident. Simply put, we learn from our mistakes, correcting our errors along the way. We all know that we have had a serious failure of the financial and investments markets due to excessive risk exposure and losses. We know that conventional statistics of normal distributions (such as used in VaR techniques) do not work when applied to accident, event and outcome trends [1, 2]. *The failure to predict is due to the failure to include the influence of human element.* We have previously shown that the dominant contribution to *all* management and system failures, outcomes and accidents is from that same inextricable and inseparable human involvement. Be they airplane, auto, train or stock market crashes, the same learning principles apply. We have shown that to quantify risk we must include the learning behaviour, quantifying outcomes rates and probabilities due to our *experience* from human decision making and involvement with modern technological and social systems, including industrial, transportation, chemical, financial and manufacturing technologies [3, 2]. These ideas and concepts include naturally not only the collective system (e.g., a bank, railway, power plant or airline) but also the individual human reliability (e.g., an investor, driver, manager or pilot).

What we know is that provided we have prior data we can now predict accurately the future outcomes rates, and define the risk exposure based on the past known and the future expected experience. That we learn from *experience* is what *all* the data show, and that experience is the past risk exposure we have all so painfully acquired as a human society. The experience measure is a surrogate for our very human risk exposure, of how long, how many, how much we have been exposed to the chance of an outcome, or to the risk of an error.

The prediction of the future rate of failures or outcomes is given from the Learning Hypothesis, being simply on the principle that humans naturally learn from their mistakes, by correcting and unlearning during and from the accumulated experience – both good and bad. The experience – however it is defined or measured – represents also not only the learning opportunity, it also is a measure of the risk exposure. The probability of error, accident, catastrophe or mistake,  $p$ , is determined by the failure rate, which derives from the number of either a successful or a failed (unsuccessful) outcome. The rate of outcomes decreases exponentially with experience, in the form of a Universal Learning Curve (ULC). Over 200 years of experience and millions of *prior*, past or historic data allow the ULC to be defined. The validation derives from massive datasets of both frequent and rare events, and now includes the following sources and outcomes, with the historical time spans shown in brackets:

- a) From official government and industry websites we analysed USA auto passenger deaths numbering 181,086 (1975-1996); 833,494 railway injuries (1975-1999); 31,770 coal mining deaths (1938-1998); some 8696 oil spills of over 1000 gallons each at sea (1961-2001); 3386 airline near misses (1980-1997), 6784 recreational boating deaths (1960-1996); and some 3386 reported mid-air near misses (1980-1997). In each case learning trends are present to some degree, and allow the learning rate “constant”,  $k \sim 3$ , to be determined.
- b) More globally, the learning data set we have amassed now contains multiple technologies worldwide: coal and gold mining injuries in Australia, UK and South Africa (1969-1999); globally over 20 million pulmonary disease deaths (1840-1970); some 284,797 cataract operations in Canada (2001-2003); 283 infant heart surgeries in UK (1984-1995); an international total of 1882 rocket launches (1962-2005); some 46,510 auto deaths (1981-1999), and 247 pilot deaths (1990-2005) in Australia; some 1652 train derailments (1988-1999) and 4568 danger signal passed (1994-2000) on UK railways; 1359 Canadian mid-air “air proximity” events (1989-1998); and notably the anti-missile interception and destruction over England of 3861 German V1 bombs (1944) in World War II.
- c) Systems that do not show significant learning, as measured by decrease or declining loss and error trends include a sample of some 376,962 deaths in reported national studies (1975-2005) of adverse medical events (errors) where work habits, work loads, traditional practices, legal issues, management and liability pressures, and patient confidentiality constraints all defy openness and error reduction. A nearly constant rate of one per thousand shipping-years pervades the thousands of global shipping losses at sea (for 1800-1950 and 1972-1997) where the continuing influence and reliance on the human element overrides massive changes in technology and the robustness of ship design. The fatal rate of some 276 commercial aircraft crashes from over 210 million flights world wide (1977 – 2000) show the lack of further learning trends, as crashes became rare events and attained the lowest rate currently known of about one per 200,000 flying experience or risk hours. In each case, large efforts are underway

nationally to understand the lack of learning trends and reduce the event rates further, for many commercial, perception and genuine safety reasons.

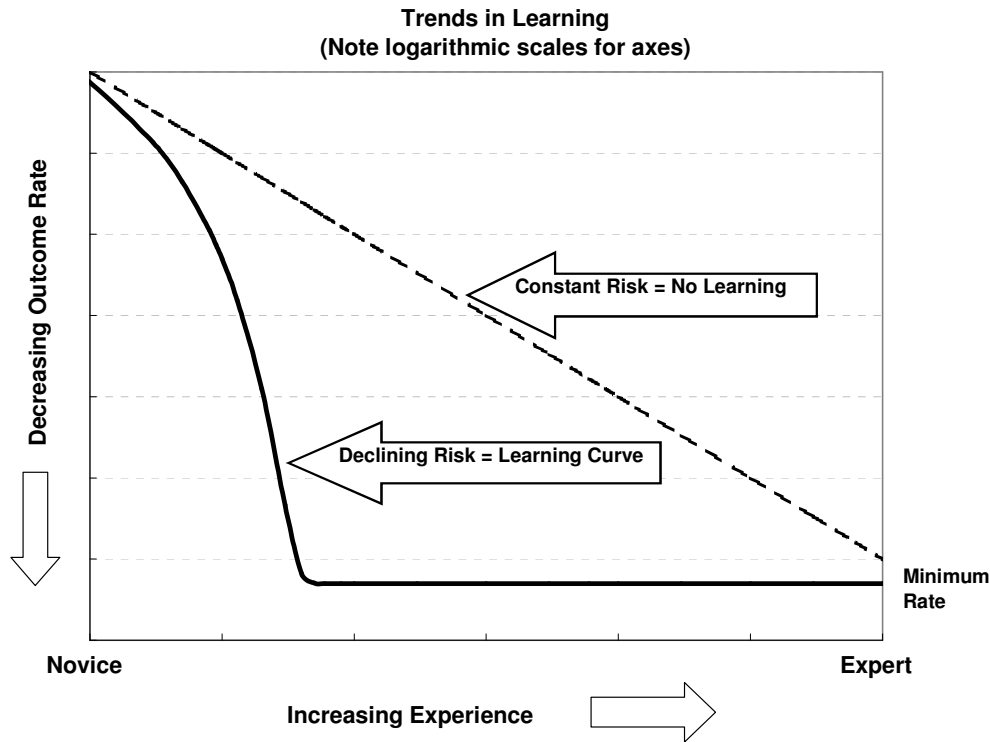
- d) For just individual actions (as opposed to system outcomes), data are available from many thousands of individual human subject task and learning trials in the psychological literature (1930-2000). These have established the rate of skill acquisition via Laws of Practice, which we have shown are consistent with the ULC, and also the reduction in response time with repeated trials. In addition, the data show that system-learning behaviour mirrors the learning trends of the individuals within. The predicted probability of operator error agrees with about 900 published French nuclear plants events (1997-1998); and with 55 or more simulator tests conducted for nuclear power plants (2003). Recovery action times and probabilities for power restoration for 148 power losses at over 100 US nuclear power plants (1986-2004) are also in agreement, as is the power blackout repair probability for New York Queen's Borough that affected (i.e., disconnected) some 175,000 customers over a period of several days.
- e) Cost data on specific unit price variations with output or sales demonstrate the learning trends causing cost reductions and productivity improvements, and include: aircraft manufacturing (1926-1936); photovoltaic cell manufacturing costs for thirty years (1968-1997); some 60 million air conditioner sales in Japan (2001); ethanol market prices changes in Brazil during process and market development (1979-1995); the declining cost of windmills for power production for Spain (1990-2001); and the UK (1992-2001 ); reduced prices for gas turbines in the USA (1990-2000); and so-called "progress curves" for manufacturing as observed for millions of units produced in factories and production lines (1969).

In all these data, we have,  $n$ , outcomes occurring in some experience,  $\tau$ . The resulting form of the learning curve is shown in Figure 1, which is a log-log plot with arbitrary units on each axis of the rate of the undesirable errors and outcomes,  $dn/d\tau$ , versus the accumulated experience, which is a surrogate for the risk exposure during actual system operation. This *risk exposure* or *experience measure* is unique for each and every system: for aircraft is the number of flights flown; for railways the train-miles travelled; for ships the shipping-years afloat; for manufacturing the number of units produced; for human errors in decision making, skill acquisition and response time it is the number of repetitive trials, etc., etc.

As we increase our experience, as individuals and systems, the event or outcome rate depends on whether, either collectively and/or individually, we follow a learning curve of decreasing risk or not, or we are somewhere in between. In Figure 1, the line labelled "learning curve" (from the Minimum Error Rate Equation) is the desirable ULC, where learning occurs to rapidly reduce the rate. This is the most likely path, and is also that of the *least* risk as we progress from being a "novice" with little experience to becoming an "expert with progressively more exposure and experience. There are no "zero defects"; there is always a finite, non-zero residual rate of error,  $\lambda_m$ , so say all the world's data. The equation that describes the learning curve is an exponential with experience:

**Failure rate,  $\lambda(\tau)$ =Minimum rate,  $\lambda_m$ +(Initial rate,  $\lambda_0$ -Minimum rate,  $\lambda_m$ )  $\times \exp(-k\tau)$**

Here in the exponential term at the end, k is the learning rate “constant”, a measure of how fast the rate decreases with increasing experience,  $\tau$ . The influence of such a learning curve is ubiquitous in economics too. If we simply replace the rate,  $\lambda$ , by the value or cost, C, the MERE turns out to be identical in form to that of the trending part of the Black-Scholes equation for portfolio cost and value, and also to the “progress curve” for cost reductions in manufacturing, with a “tail” of non-zero value that corresponds to the minimum possibly achievable in any competitive market system. The difference is that in these cases the experience parameter,  $\tau$ , is conventionally taken as either time (for stock or equity values variation) or accumulated units manufactured (for production prices changes), and a key question is what to adopt for the relevant experience measure.



**Figure 1 The ULC and Constant Risk Lines**

Since Figure 1 is a log-log plot (scale units are factors of ten on each axis), any line of constant risk is then a straight line of slope minus one, where the event rate,  $\lambda$ , times experience,  $\tau$ , is the constant number of events, n. Hence,  $\lambda = n/\tau$ , and for the first or rare event, n = 1, which is the “constant risk” dashed line shown in Figure 1. The rate decreases inversely with the risk exposure or experience, so importantly, at little of no experience, the initial rate is given by  $\lambda_0 = 1/\tau$ , which is exactly the form of the rare events as derived from commercial aircraft crashes. Since the rate falls simply as the inverse of the experience,  $1/\tau$ , it exhibits no learning trend and simply represents when random or “rare” events happen. As we shall see this risk path is the initial rate and also

the “fat tail” that worries and confounds conventional risk and value analysts. We call this prediction a White Elephant when it underestimates the risk.

In terms of probabilities as a measure of risk, instead of rates, the above equation can be integrated to yield an expression that in words implies:

Risk exposure probability is due to the minimum risk plus the initial risk exposure less the reduction in risk due to learning.

For any real, not hypothetical system the minimum achievable failure rate does not appear to change and has not changed for over 200 years, depending solely on our experience and risk exposure measure for a given system. So conversely, the *systemic risk is dependent on the risk exposure measure*.

### **The Seven Commonalities of Rare and Terrible Events: Risk Ratios**

What do large disasters in financial systems like the Great Crash of 2008 [4] have in common with the other major events? These have happened in multiple technologies and industries, such as in industries as diverse as aerospace (Columbia and Challenger Shuttle losses) [5], nuclear (Davis-Besse plant vessel corrosion) [6], oil (Texas City refinery explosion) [7], chemical (Toulouse ammonia plant explosion) [8] and transportation (the Quebec overpass collapse)[9]. The common features, or as we may call them the Seven Themes, cover the aspects of causation, rationalization, retribution, and prevention that are all too familiar:

*First*, these major losses, failures and outcomes all share the same very same and very human Four Phases or warning signs: the unfolding of the precursors and initiating circumstances; the confluence of events and circumstances in unexpected ways; the escalation where the unrecognized unknowingly happens; and, afterwards, denial and blame shift before final acceptance.

*Second*, as always, these incidents all involved humans, were not expected but clearly understandable as due to management emphasis on production and profit rather than safety and risk, from gaps in the operating and management requirements, and from lax inspection and inadequate regulations.

*Third*, these events have all caused a spate of media coverage, retroactive soul-searching, “culture” studies and surveys, regulation review, revisions to laws, guidelines and procedures, new limits and reporting legislation, which all echo perfectly the present emphasis on limits to the “bonus culture” and “risk taking” that are or were endemic in certain financial circles.

*Fourth*, the failures were so-called “rare events” and involved obvious dynamic human lapses and errors, and as such do not follow the usual statistical rules and laws that govern large quasi-static samples, or the multitudinous outcome distributions (like normal, lognormal and Weibull) that dominate conventional statistical thinking, but

clearly require analysis and understanding of the role of human learning, experience and skill in making mistakes and taking decisions.

*Fifth*, these events all involve humans operating inside and/or with a system, and contain real information about what we know about what we do not know, being the unexpected, the unknown, the rare and low occurrence rate events, with large consequences and highlighting our own inadequate predictive capability, so that to predict we must use Bayesian-type likelihood estimation.

*Sixth*, there is the learning paradox, that if we do not learn we have more risk, but to learn perversely we must have the very events we seek to avoid, which also have a large and finite risk of re-occurrence; and we ultimately have more risk from events we have not had the chance to learn about, being the unknown, rare or unexpected.

*Seventh*, these events were all preventable but only afterwards – with 20/20 hindsight soul-searching and sometimes massive inquiries reveal what was so obvious time after time; the same human fallibilities, performance lapses, supervisory and inspections gaps, bad habits, inadequate rules and legislation, management failures, and risk taking behaviours that all should have been and were self-evident, and were uncorrected.

All of these aspects were evident in the financial failure of 2008, in the collapse of major financial institutions and banks. These rare events are worth examining further as to their repeat frequency and markets probability: recessions have happened before but 2008 was supposedly somewhat different, as it was reportedly due to unbridled systemic risk, and uncontrolled systemic failure in credit and real estate sectors. This failure of risk management in financial markets led to the analysis that follows, extending the observations, new thinking and methods developed for understanding other technological systems to the prediction and management of so-called “systemic risk” in financial markets and transactions. We treat and analyze these financial entities as “systems” which function and “behave” by learning from experience just like any other system, where we observe the external outcomes and failures due to the unobserved internal activities, management decisions, errors and risks taken.

The past outcome data provide the past failure rate. To determine the future risk, we must distinguish between the past (statistically, the known prior) and the future (statistically, the unknown posterior). So what does the past tell us about the future? To predict an outcome, any event, we must go beyond what we know, the prior knowledge. Somehow we have to project ourselves into an unknown future, with some measure of confidence and uncertainty, based on both our rational thoughts and our irrational fears, using what we know about what we do not know. This leads us into the somewhat controversial arena of prediction using statistical reasoning, a subject addressed in great detail elsewhere [10],

The *conditional* future is dependent, albeit with uncertainty, on the past, as per Bayes reasoning [10, 11]. The probability or chance of an unknown event is dependent on something called the *likelihood*, which itself is uncertain but provides a rational

framework for projection. The likelihood itself is inversely dependent on the prior number of outcomes, and if there are none so far, we just have the Bayesian failure rate of the past based on our (known) experience to date.

The Likelihood formally adjusts the past, prior or known probability and produces the future or Posterior probability. So conditionally dependent on what we already know we know has already happened in the past, according to the thinking of the Reverend Thomas Bayes (1763) and of Edwin Jaynes' (2004) rigorous analysis:

Future chance (posterior probability,  $p(P)$ ) = Past or prior probability,  $p$ , times Likelihood

The Likelihood multiplier,  $p(L)$ , whatever it is and however derived (by physical argument, guess, judgment, evidence, probabilistic reasoning, mathematical rigour or data analysis) is the conditioning factor which always alters the past whatever and however it is estimated. Even if the past was indeed "normal" the likelihood can even change the future to include rare events and unknown unknowns.

The risk ratio (RR) can then be defined as ratio of the future posterior probability,  $p(P)$ , of an adverse event (accident, outcome, error, or failure in the future) to some known past or present failure probability,  $p(\tau)$ , based on the prior accumulated experience, as a function of the future risk exposure or experience, or

$$RR = p(P)/p(\tau).$$

From the above Bayesian equation this is equivalent to defining the Likelihood,  $p(L)$ , where for low probabilities or rare events the posterior,  $p(P)$ , itself is numerically very nearly equal to the rate of events, or the failure rate,  $p(P) \sim f(\tau) \sim \lambda$ . This result follows directly from the so-called "generalized Bayes formula" [12, 2] that defines the Likelihood as the ratio of the probability of outcomes occurring in the next experience interval to the probability that outcomes have already occurred during the past experience.

So for low probability events, outcomes or disasters ( $p(\tau) \ll 1$ ), the Risk Ratio becomes simply the future predicted by the past since:

$$RR = p(P)(1-p(\tau))/p(\tau) \sim p(P)/p(\tau) \sim \lambda(\tau)/p(\tau)$$

which is the ratio of the known past rate and prior probability.

We show the Risk Ratio, RR, prediction for rare events with little learning ( $k \sim 0.0001$ ) in Figure 2 versus a series of curves ( $k$  from 0.1 – 0.001) representing slow to negligible learning, where the Risk Ratio clearly has a slope varying as,  $1/\tau$ . The key observation is that the future risk predicted by the risk ratio, RR, still does not fall much below  $\sim 10^{-5}$  at large risk exposure, which corresponds to the plateau, or "fat tail", caused by the lowest attainable but finite and non-zero failure rate that is observed for any system anywhere and everywhere in the world.



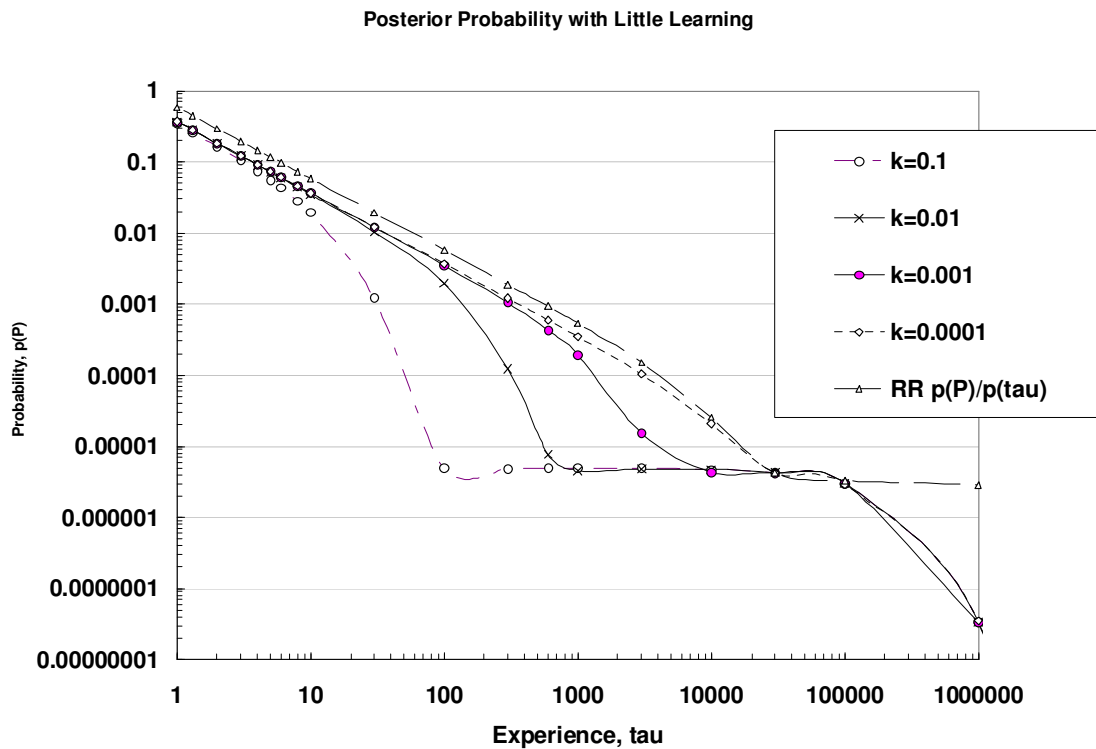


Figure 2 Comparisons of the Risk Ratio Predictions

So what then is the resulting Posterior probability,  $p(P)$  in the future? It is shown in Figure 2 for a series of cases with varying learning or knowledge acquisition from increasing risk exposure or accumulated experience. These cases are represented by the range of values shown for the learning “constant”,  $k$ , where progressively lower values mean less and less learning. As can be seen, if learning is negligible so,  $k$ , is very small (say, 0.0001) then the event probability decreases almost as a straight line of constant risk,  $1/\tau$ , as it should; for larger  $k$  values a distinct kink or plateau occurs due to the presence of the always finite, non-zero failure rate due to the human involvement.

### Predicting Rare Events: Black Swans and White Elephants

Colloquially, a black swan is an unexpected and/or rare event, one that dramatically changes prior thinking and expectations.

Because rare events do not happen often, they are also widely misunderstood. Perhaps even previously unobserved, they are called “unknown unknowns” [13], or “Black Swans” [1] precisely because they do not follow the same “rules” when having many or frequent events. Think of the space shuttle crashes we have already seen, the global collapse of financial companies that have occurred, or an aircraft apparently falling from the sky as it did recently over the Atlantic. They are the things we may or may not have seen before, but certainly did not expect to happen. So when they do happen, perhaps

even when being thought not possible, they do not apparently follow the trends, expectations, rules or knowledge we have built up for more frequent happenings. There is no assured, easy or obvious “alarm”, indicator or built in warning signal, derivable by adjusting “filters” or data smoothing techniques. As noted in [14], “*Whether these alarms are deemed informative depends on their association with subsequent busts. The choice of a threshold above which an alarm is raised presents an important trade-off between the desire for some warning of an impending bust and the costs associated with a false alarm. Nonetheless, even the best indicator failed to raise an alarm one to three years ahead of roughly one-half of all busts since 1985. Thus, asset price busts are difficult to predict.*” This is a 50% or even chance, which are no better odds than just tossing a coin.

In statistical language and usage, the rare events do not follow or fit in with the usual distributions of previous or expected occurrences. The frequency and/or probability of occurrence lies somewhere outside the usual many expected multiples of the standard deviation for any sample distribution. We may not even have a distribution of prior data anyway. In fact, Taleb [1] spends a considerable part of his popular book “The Black Swan” discussing, discounting and dismissing the use of so-called “normal distributions” such as the Gaussian or bell-shaped curves simply because they do not and cannot account for rare events even though many humans may think that they do. Also rare events, like all events, as we have said, are always due to some apparently unforeseen combination of circumstance, conditions, and combination of things that we did not foresee, and all include the errors in our human made and managed systems (the Seven Themes).

By citing many empirical cases, Taleb [1] also further argues forcibly that this “scale” variation destroys any and all credibility of using any Gaussian or “normal” distribution for *prediction*. In that limited sense, he is right, as conventional sampling statistics based on fitting to some “normal” distributions using many observations is totally inapplicable for low probability, one-of-a-kind rare, so-far-unobserved or unknown events. To make a true prediction we must still use what we know about what we do not know, and we now know that the *relevant “scale” is in fact our experience or risk exposure*, which is what we have anyway, and is the basis for what we know or do not know about everything.

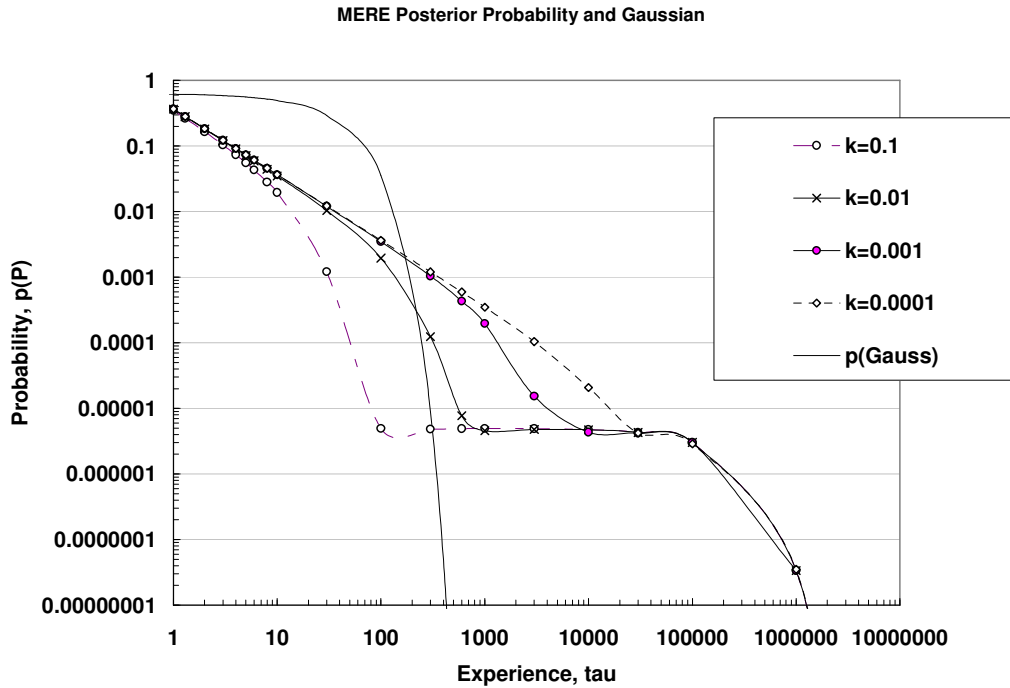
In Figure 3, we show the one-on-one head-to-head comparison of a normal (Gaussian) bell-shaped distribution<sup>1</sup>, compared to the reality of learning variations as they affect probability: it is clear that the Gaussian or normal distribution seriously underestimates risk, in this case the probability of an outcome, for large experience. This inability of standard methods to predict the extrema of the distributions is itself well known- but less well known is that the probability increase or plateau is due to the human element.

So the future chance, or posterior, of any event, even of an unknown unknown, *is* in fact given by estimating the Likelihood,  $p(L)$ , something Taleb does not discuss at all.

---

<sup>1</sup> The Gaussian (or normal) distribution shown in Figure 3 is  $p(P) = 23\exp(-0.5(\tau+290)/109)^2$ , and was fitted to the MERE learning curve using the commercial statistical software routine TableCurve 2D.

Instead, the concept of “scalability” was invoked, which we have now shown and will demonstrate is actually the same thing as a conditional probability of whether it will occur, but disguised as another White Elephant.



**Figure 3 Predictions: illustrating the Gaussian distribution failure to include the "fat tail" due to the influence of the human element**

The impact of rare events can vary, particularly because they were somehow disruptive, unexpected or not predicted. So impacts can be large, as for a financial crisis that affects everyone’s credit or bank account, [1]; or they can be negligible because they do not affect the overall industry but only the participants, as for a commercial airplane crash. But both do not happen very often. Because events occur randomly, we find it difficult to predict when and where they will happen, and can do so only with uncertainty. So with rare events we are more uncertain as we have had limited learning opportunity, and we fear the unknown. The risk we determine or sense can be defined as the uncertainty in the chance of such an event happening. It is perceived by us, individually and collectively, as being a high risk or not based on how we feel about it, and have been taught, trained, experienced, learnt, or indoctrinated. The randomness is then inherent in the learning processes, in the myriad of learnt and unlearnt patterns, neural firings, legal rules, acquired skills, written procedures, unconscious decisions, and conscious interactions that any and all humans have in any and all systems. Perversely, only by having such randomness, learning, skill, trial and error can order and learning patterns emerge. We create order from disorder, learning as we go from experience and risk exposure, discerning the right and unlearning the wrong behaviours and skills. So a rare Black Swan of major impact is indeed a White Elephant of no intrinsic value unless and only if we are learning.

We need to know what we do not know. We cannot know what happens inside our brains and see the how the gazillions of neural patterns, pathways and possibilities are wired, learnt, interconnected, rationalized and unlearned. We cannot know the millions of things that any group of people will talk about, learn, exchange, review, revise, argue, debate, reject, use and abuse, each and every day, 24/7. We cannot know all about how a machine or system will behave when subjected to the whims of inadequate design, poor maintenance, extreme failure modes, external damage, and poor or unsafe operation. What we do know is that, because *we are human*, we do learn from our mistakes: this is the Learning Hypothesis [15, 16, 3]. The rate at which we make errors, produce outcomes, and cause events reduces both as we gain experience and if and as we learn. We make mistakes because we are human: the fat tail, the rare event, is because we are human. If and as we gain experience, this is equivalent to increasing our risk exposure too. The risk increases whether by driving on the road, by trading stocks and investments, or by building and operating a technological system like a ship, train, rocket or aircraft.

Consistent with the principles of natural selection, those who do not learn, those who do not adapt and survive, are the failures and extinctions of history, overtaken by the unexpected and mistakes, the errors and the black swans *of the past*.

### **Failure to Predict Failure: Scaling Laws and Risk**

What do we know about what we do not know? We know that the four categories of knowns and unknowns are the Rumsfeld quartet:

Known knowns – what is expected and already observed (in the past)

Known unknowns – unexpected but observed outcomes (past outcomes)

Unknown knowns – expected and not yet observed (in the future)

Unknown unknowns – unexpected and not yet observed (future outcomes or rare events)

This is analogous to drawing both outcomes and non-outcomes from Bernoulli's urn [2], and the probability of a rare (unknown) event is determined if all we do is assume that it exists. Thus, we have turned a Black Swan into a White Elephant – the fact that we have not observed it, do not know if it exists, but can rationally discuss it allows the fear, dread and risk perception to be quantified. This is precisely what Taleb recommends – taking precautions against what it is we do not know about what we do now know.

We have defined a risk ratio, RR, which depends on the prior failure rate. But for a rare or unknown event the posterior probability of an unknown unknown,  $p(U,U)$  which has not happened yet is finite and is given by analogy to the “case of zero failures” [17]. We can then obtain the estimate for knowing the posterior (future) probability of the unknowable as [2]:

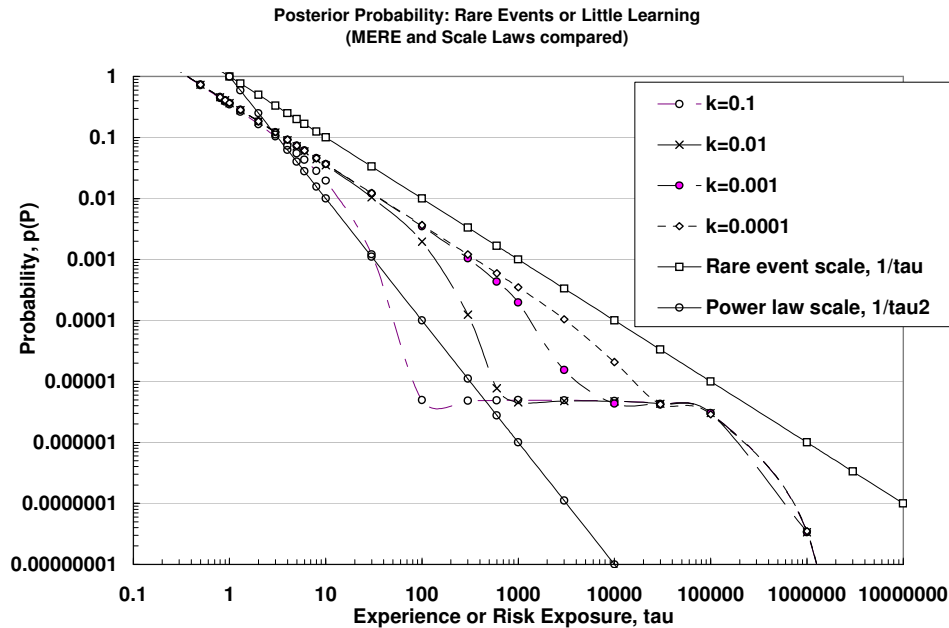
$$P(U,U) \sim (1/U\tau^2) \exp-U,$$

where, U, is some constant of proportionality. This order of magnitude estimate shows a clear trend of the probability decreasing with increasing experience as an inverse square

power law,  $\tau^{-2}$ . For every factor of ten increase in experience measured in some tau units,  $\tau$ , the posterior probability falls by one hundred times. It does not matter if we do not know the exact numbers: the trend is the key for decision making and risk taking. *The rational choice and implication is to trust experience and not to be afraid of the perceived Unknown.*

The risk of an unknown unknown therefore decreases with our increasing experience, or risk exposure. So the White Elephant is *precisely* the case of little or no learning corresponding *exactly* to a scaled probability inverse law, i.e.,  $p(P) = n/\tau$ , where the number of events,  $n$ , is one ( $n=1$ ), simply because it is that first and rare event that was never previously observed or known. So the probability,  $p$ , of any single rare event is always,  $1/\tau$ , the inverse of (one divided by) the exposure or experience measure, or “scale”. As shown before in Figure 2, this is also a measure of the Risk Ratio, RR, and is equivalent numerically to the failure rate,  $\lambda$ . So also shown in Figure 4 are the so-called “scalable” or pure “power” laws discussed by Taleb [1], where the probability is *assumed* to fall as the more general inverse power law,  $p(P) = 1/\tau^\alpha$ .

Corresponding to the prior and the posterior variations without significant learning, for illustration,  $\alpha$ , is often taken as lying in the range between 1 and 2 which assumed values nicely cover the “true” curves for novice or zero experience, varying as,  $1/\tau$ , and for “unknown unknowns” varying as,  $1/\tau^2$ . But these are “constant risk” lines that do not give the detailed shape or slope variations since they do *not* reflect the learning opportunity and the finite non-zero risk rate. Basically the incorrect inexorable decrease in risk predicted by a scale law is offset by the inevitability of risk due to the human element, causing the “fat tail” or plateau in the probability graph. At a future (posterior) probability of order  $p < 10^{-5}$  the line intersects the learning curves, the Black Swan truly becomes a White Elephant, being of less value or lower risk than the actual.



**Figure 4 The Rare Event Prediction**

Popular because of its simplicity, the inapplicable power law form is widely used in the field of economics (known as an “elasticity”) when fitting the exponent to price or response time reduction [18]; in cognitive psychology (known as a “law of practice”) when applied to trials that constitute repetitive learning [19]; and in damage estimation for industrial failures and collapses [20]. The exponent is roughly constant only over limited ranges of data, otherwise it fails in extrapolating magnitude or trend [2]. In fact in statistics, this form of inverse “power law” type of relation is often known as a Pareto distribution<sup>2</sup>, and Woo [21] explicitly cautions that: “*parametrizing a natural hazard loss curve cannot be reliably reduced to a statistical analysis of loss data, e.g., fitting a Pareto curve: damaging events are too infrequent for this to be sound.*”

In fact, this failure to predict may even explain the proven poor capability of many economic models that by using a constant “elasticity” between price and demand and extrapolating we now know from data do not predict well! This general “power law” form also fits social trends, such as word usage, books sold, website hits, telephone calls, and city populations, leading Taleb [1] to further argue that this form represents true “scalability” which we can now recognise as the *fundamental* connection to learning and risk exposure. Arbitrary adjustment of the exponent,  $\alpha$ , in economics, social science and cognitive psychology is an attempt to actually account for and fit what we observe, but without trying to understand why the exponent is not unity nor placing limits on the extrapolations made beyond the known data used for the original fits. We now know and can see from Figure 4 that the exponent variation *in reality* is due to the presence and

<sup>2</sup> Also termed the hyperbolic or power-law distribution, the form given by Woo for natural catastrophes is:  $p(\tau) = ba^b/\tau^{b+1}$ , where  $a$  and  $b$  are constants, the so-called “location” and “shape” parameters.

effects of *learning*, with the larger exponent values and steeper slope encompassing the variation between the learning curves (Figure 3).

This presence of learning effects explains nicely the actual range of empirical values for the exponent,  $\alpha$ , quoted by Taleb and others of between 1 and 2 - *some systems evidently exhibit more or less initial learning than others* as is shown in Figure 4. In the inverse power law simplification, if there are no events by definition there is and can be no learning. Strictly we know this is not true, as we also learn something from the many and often irritating non-events and near misses. This so-called incidental learning leads to the other extreme case of “perfect learning” [2], where the event outcome probability still follows a learning curve until we have just one event, and then subsequently plummets to zero.

*We stress here, in italics, that the power law form is a natural, simplified limiting variant of the more general “learning curve”, which naturally then also encompasses the occurrence of rare events.*

The analysis of risk ratios due to the financial cost of individual events assumes that big losses or damage occur less often i.e., are rare or lower in frequency. For example, Hanayasu and Sekine [20] argue that the rate of financial “damage” of events in industry decreases with the inverse of the damage or loss. So generally the frequency of an event decreases with increasing cost as the probability density,

$$dp/dt \approx \text{constant} / h^{q+1}$$

Here,  $q$ , is yet another power law exponent chosen to fit some damage data, and is always such that  $q > 1$  so Hanayasu and Sekine assume that it lies in the range  $2 < q < 3$ . When the slope is an inverse cube such that,  $\alpha \sim 3$ , there is a very rapid decline. We analyzed this approach [2] and found the risk ratio, RR, or damage ratio referenced to some initial known value,  $h_0$ , and probability,  $p_0$ , is then given by:

$$RR = (h/h_0) = (p/p_0)^{1/q}$$

Extrapolation of the fitted line beyond the data range given shows a much faster decrease in risk ratio than usually observed, or expected from a learning curve with a finite minimum that flattens out. So the basic problem is that extrapolation of the size of the loss according to this “power law” (although it is not really a law at all) produces inaccuracy outside the known data range, does not account for learning, and also does not allow for the finite non-zero contribution of the human element (the extra “fat tail” shown in Figure 2). We have fitted a MERE curve also to these damage data, and as a result the forward risk exposure or uncertainty is grossly *underestimated* because of omitting the human learning element. This is really uncertainty: we are predicting the variation in how big the losses will be for unknown events, based on what we know.

The chance of an unknown unknown or rare event also depends on whether or not you learn! Conversely, rare events and Black Swans are also simply events for which we have

little or no learning. The argument is then wrong that this type of inverse power variation represents “true” randomness, where there is no pattern other than that which is “scale” invariant (like fractals). In fact the variation in probability or risk in reality is all due to whether we have been learning or not, at what rate we make or have made mistakes both in the past and in the future. The true natural “scale” for all human-based systemic risk we have shown repeatedly is our *experience*, however that is defined and accumulated, as learning is not invariant with risk exposure. What we know about the unknown is that we are human and remain so, learning as we go.

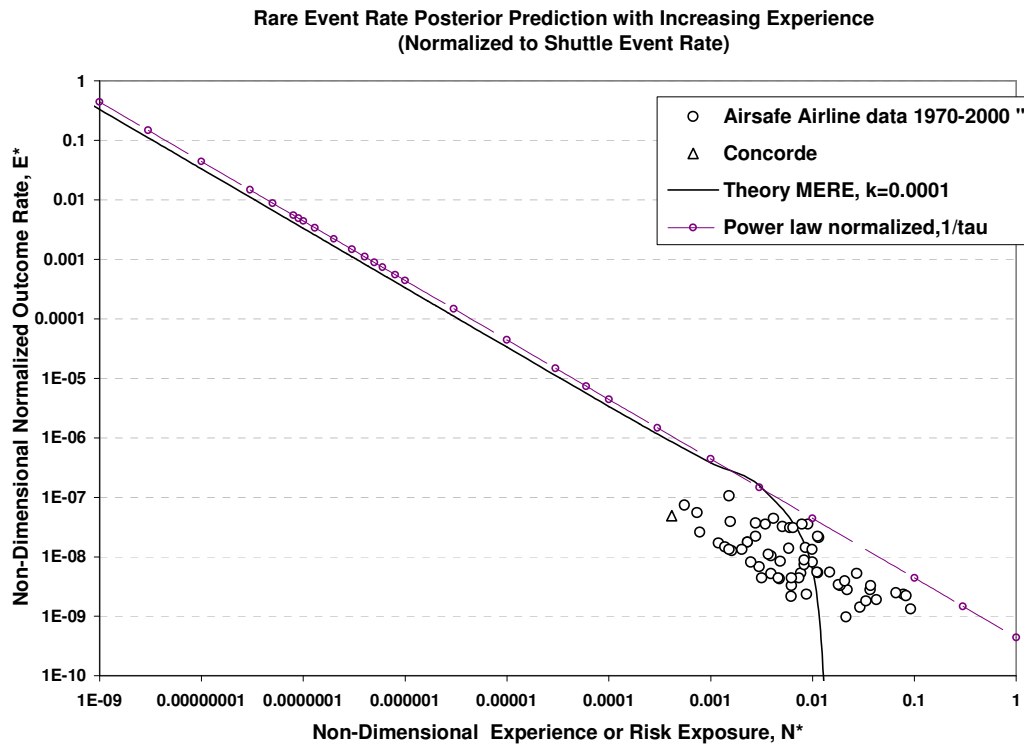
For the future unknown experience, the average future failure rate,  $\langle \lambda \rangle$ , we will observe over any future risk exposure or operating interval,  $\tau - \tau_0$ , is obtained by averaging the varying failure rate over that same observation or risk exposure interval, so:

$$\langle \lambda \rangle = \frac{1}{(\tau - \tau_0)} \int_{\tau_0}^{\tau} \lambda(\tau) d\tau$$

Clearly, the *apparent* average rate also depends on the interval,  $\tau - \tau_0$ , over which we start and finish observing, or choose to record outcomes, or happen to be present, or are risk exposed.

We can show how these ideas work in practice by comparing to actual data for rare events, although this is strictly an oxymoron, as if the outcomes occur they are no longer “rare” or become known unknowns. The data available is the case we have analyzed in detail before [2, 3] for fatal commercial airline crashes between 1970-2000. The case is relevant as the airline industry is regarded as relatively safe, and having perhaps attained the lowest possible event rate. Over this 30-year period using modern jets, some 114 commercial passenger airlines accumulated about 220 million flights, and there were about 270 fatal crashes, excluding hull losses (plane write offs) with no deaths.





**Figure 5 The prediction for rare aircraft crashes**

The analysis shows that the airlines having the least experience have the highest rate per flight, the airlines overall having descended the learning curve and achieved their lowest possible “rare” crash rate. So for this case, flights accumulated represent a convenient measure of the risk exposure and learning “scale”. As we would expect, as seen in Figure 5, there is therefore an almost constant risk where the fatal crash rate indeed varies inversely as,  $\lambda \sim 1/\tau$ , the number of accumulated flights being the measure of both the learning experience and risk exposure<sup>3</sup>. For reference, the maximum interval (but not the average) achieved by commercial airlines is about one fatal crash every 100,000 to 200,000 flying-hours as the measure of risk exposure, and is typical of the lowest failure rate achieved by technical systems with human involvement [3]. The only larger interval found is for systems like dams, where humans are passive and not actively and/or continuously involved in the day-to-day system performance and operation.

But the relative future risk of a mature technology, as measured by the non-dimensional posterior outcome rate, is negligible compared to that for new technology. The plunge in the *future* prediction,  $p(P)$ , of the risk at large experience, or the “thin tail” appearing in the end of the fat tail, is due to the *prior* probability becoming nearer and nearer to certainty ( $p \rightarrow 1$ ) at large enough experience or risk exposure since the failure rate (according to all the world’s known outcome data) is never, ever zero. Thus, we have

<sup>3</sup> The fundamental problem and seeming paradox with using event rate as a measure of risk for rare events is that the rate and number seemingly fall with increasing experience (not just time), giving an apparent decrease when in fact the risk of a random outcome is effectively still constant.

found a basis on which to make predictions of all such rare unknown unknowns, based on the (equally) rare prior outcomes from many disparate sources.

We have already recently used the methods and ideas discussed here and in our book [2] to risk, failure rate and reliability prediction for many important cases. These include human errors and recovery actions in nuclear power plants [22]; predicting rocket launch failures and space crew safety for new systems [23]; the time it takes for restoration of power following grid failure (or “blackout”) [24]; predicting the rate of failure of heat exchanger tubing in new designs [25]; and the quantitative tracking of learning trends (“safety culture”) in management and operation of large offshore oil and gas facilities [26]; and about 30 other key examples. While each case has its own fascinating and unique experience and data, all examples and cases can be reduced to the common learning basis, and all follow the universal laws and rules for the outcomes due to us, as humans, functioning in modern society and technological systems, whether we know it or not.

### **The Financial Learning trends: economic growth rates, failure and stability**

Like other systems with failures and outcomes, there are a lot of financial system data out there, both nationally and globally, and data are key to our understanding and analysis. What are the right measures for failure (errors) and experience in financial systems? Can the market collapse be predicted using these measures? As an exercise in examining those questions, we explored the publicly available global financial data from the World Bank and the IMF, covering the years up to the Great Crash or “bust” of 2008. This was widely attributed to the failure of the credit markets, due to the collateralizing of risky (real estate) debt assets as securities in the developed economies and financial markets. The present analysis is to determine the presence or not of precursors, the evidence or not of learning trends, and prediction of the probability of failure using the prior data.

Let us make a financial market system prediction based solely on what we know about other system failures. According to the data (and as shown in Figures 2, 3 and 4), we have learnt that there is an apparent fundamental and inherent inability, due to the inseparable involvement of humans in and with the technological systems, for the posterior (future) probability of an outcome to occur with a probability of less than  $p(P) < 10^{-5}$ . This corresponds to the lowest observed rate of one outcome or failure in about 100,000 to 200,000 experience or risk exposure units [2, 3]. If the global financial “market”, including real estate equities and stocks, is now defined as the relevant *system with human involvement*, and a trading or business experience of 24/7/365 taken as the appropriate risk exposure or experience measure, this implies we may *expect and predict an average “market failure” rate* ranging from not less than about once every ten years and not more than every twenty years. If lack of economic (GWP and/or GDP) growth, with financial credit and market collapse is taken as a surrogate measure of an outcome or failure<sup>4</sup>, there has been apparently four relatively recent “crises” in the World (in about

---

<sup>4</sup> The recent IMF World Economic Outlook 2009 in fact shows for the 2008 crisis there is a relation between household liabilities and credit growth in relation to GDP growth ([14, Figure 3.10]).

1981-2, 1992-3, 1997-8 and 2008-9), and five “recessions” in the USA (circa 1972, 1980, 1982, 1990, and 2008) in the forty-year interval 1970–2010 [4], being an *average* risk interval of between eight (nationally) to ten (globally) years. In fact, in the full interval of 1870-2008, the IMF listed eight globally significant financial crises in those 138 years (the above four listed plus 1873, 1891-1892, 1907-1908, 1929-1931), or ten when including the two World Wars [see 4, Figure 4.1]. All these various crises give an *average* interval of about one failure somewhere between every 8 to 17 years, an agreement surprisingly close to and certainly within our present predictive uncertainty range of one about every ten to twenty years.

This present purely “rare event” prediction is a result that was not anticipated beforehand, and is based on failure data from other global and national *non-financial* systems, implying that the very same and very human forces are at work in financial systems due to human fallibility and mistakes. The present rate-of-failure approach contrasts squarely with many other unsuccessful predictive measures [27], and short and long-term bond rate spreads using “probit” probability curves tuned to the market statistical variations [28]. So although we cannot yet predict exactly when, we can now say that the “economic market place” (EMP) is behaving and failing *on average* in the same manner and rates as all other known homo-technological systems. We presume for the moment that this is not just a coincidence, and that the prior historical data are indeed telling us something about the commonality and causes of random and rare fiscal failures, and our ability or inability to predict systemic risk. So we can now seek new measures for predictors or precursors of market failure and stability based on what we know.

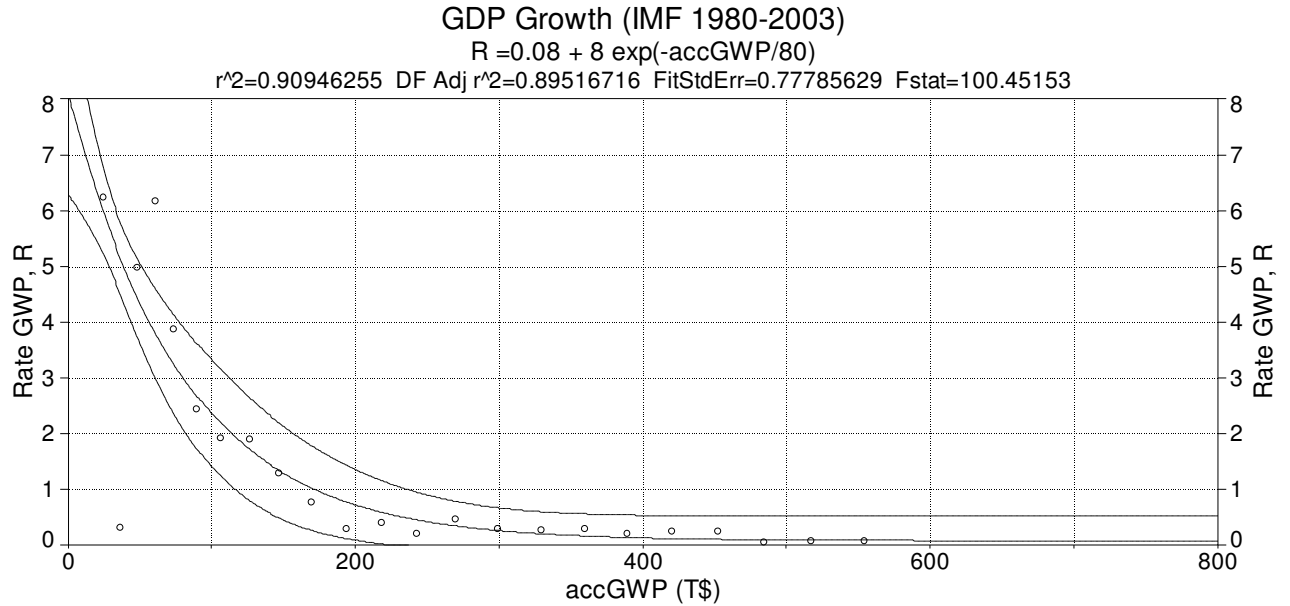
We already know that the chance of such a major event “ever happening again” is given by the matching probability using conventional statistics, and this has the value of ~0.63, or about an equal chance of happening or not [2]. This is a sure Repeat Event Prediction (REP) of a nearly equal chance. So for managing risk, we should expect another collapse based solely on this analysis, and probably with about the same *average* 10 to 20-year interval unless some change is made that impacts the human contribution. The inevitability of failure is rather disheartening, and although uncomfortable seems to be the reality, so we should all at least proactively plan for it and hence be able to manage and survive the outcome, which is risk mitigation.

Having established the possible relevance of GWP and GDP, as an initial step the measure of the outcome rate is taken to be the % growth in GWP and GDP (positive growth being success, negative growth being failure), and the relevant measure for experience and risk exposure for the global financial system as the gross world product, GWP (T\$), not in the usual calendar years as the interval over which the data are usually presented.

The result of the ULC analysis is shown in Figure 6 for the interval 1980-2003 [14], where the GDP growth rate, R, is the MERE learning curve form:

$$R, \% \text{ GWP} = 0.08 + 8 \exp\text{-accGWP}/80$$

The success rate is decreasing exponentially, and this expression is correlated with the data to an  $r^2 = 0.9$ , and importantly shows that by a GWP of order \$600T the overall global growth rate is negligible ( $<0.1\%$ ).



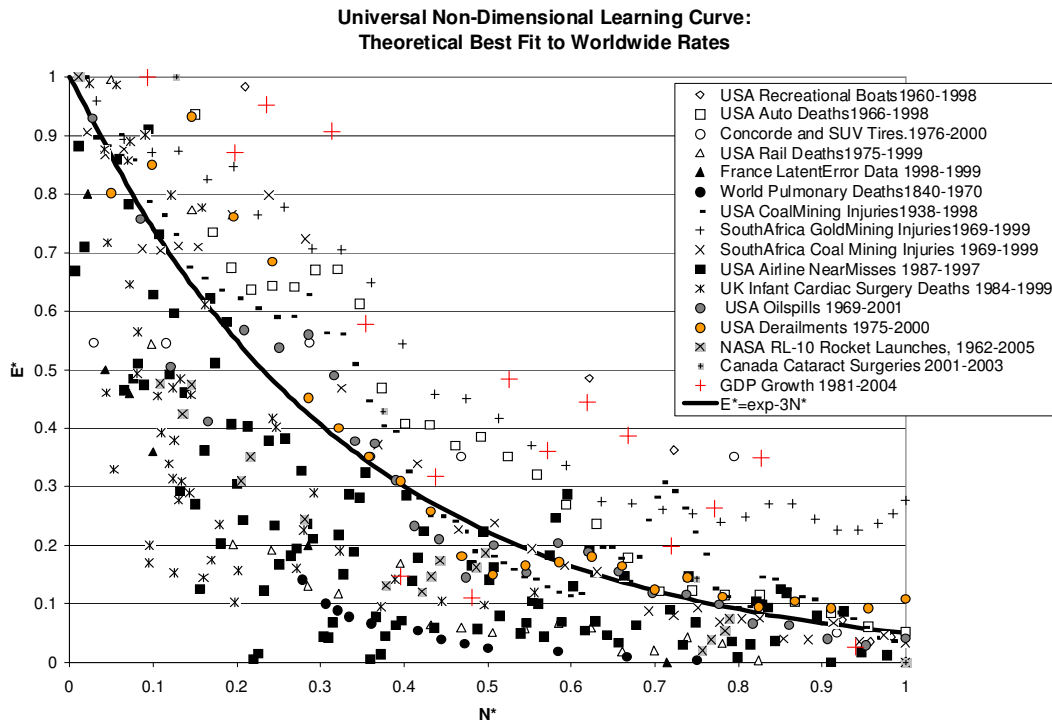
**Figure 6 The GWP growth rate curve**

It is worth noting that, as might be expected in global trading, the magnitude and growth many economies are apparently highly correlated with the accumulated GWP, so will follow similar trends as we see later. For example, the straight line that gives the relation between the USA GDP and the GWP for the interval 1981 –2004 is:

$$\text{GDP (USA, \$B)} = 15\{\text{accGWP}(\$T)\} + 3210,$$

with a correlation coefficient of  $r^2 = 0.99$ . The magnitudes are hence very tightly coupled; but here we do not have to decide which is cause and which is effect (i.e., is the change in one due to the other, or vice versa?)

As a comparison to the learning trends given by the ULC, we also plotted these growth decline data (shown as large crosses) in non-dimensional form against all the other world outcome data with the result shown in Figure 7. The trend is similar and within the scatter: basically the rate of growth, being defined as failure when declining, is tracking other adverse outcome data.

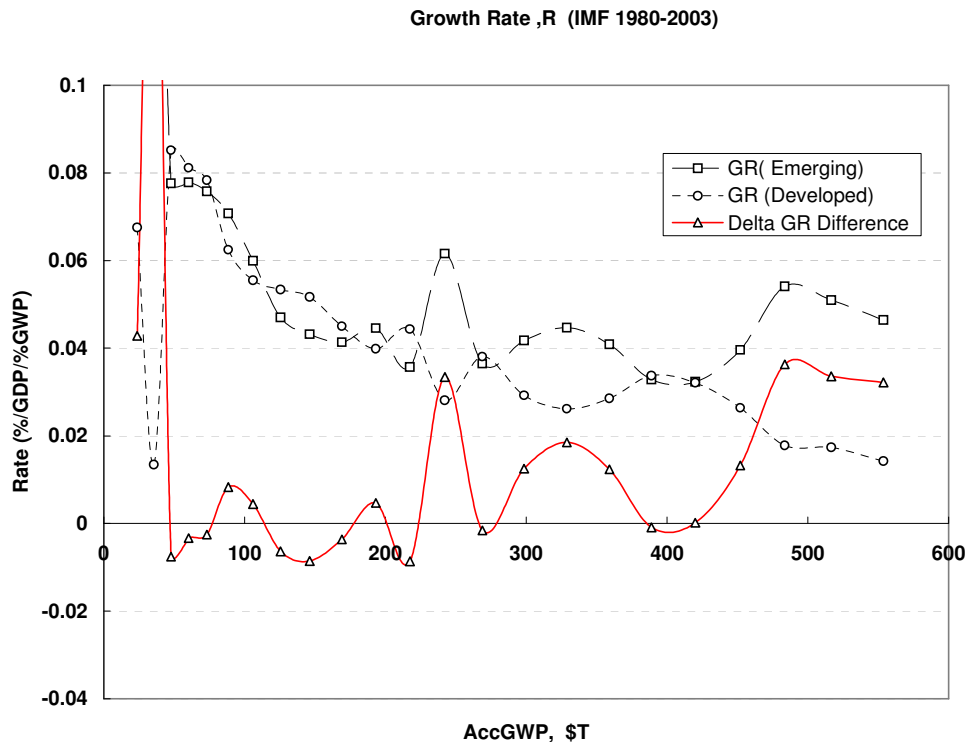


The implication is intriguing: if declining growth rates are equivalent to an error, then the developed economies suffering declines in growth had even “learnt” to further reduce their growth. To compare growth rates, the IMF and World Bank have also separated out the percentage GDP growth rates for “emerging” or developing countries/economies from “developed” or “advanced” countries/economies [27].

Now the percentage growths are based on very different totals, so just for a comparison exercise, the % growth rate,  $\pi\text{GR}$ , in each grouping was defined relative to the absolute growth in the world, or GWP, as:

$$\pi\text{GR} (\%/\$T) = \% \text{ GDP Growth} / (\text{GWP } \$T \times \text{World } \% \text{ Growth})$$

In effect, this is a measure of the rate of economic growth rate relative to the total available economic growth “pie”. The relative growth rate data calculated in this manner for the two groupings are shown in Figure 8 as a function still of the accumulated GWP, as well as the delta (or difference) in relative growth rate,  $\pi\text{GR} (\text{developed}) - \pi\text{GR} (\text{emerging})$ , between the two “types” of economies. The reason for taking the *accumulated* GWP as the experience measure is this is presumably some measure of available learning experience and risk exposure in the global trade between the two groups, and of the total available “pie”.



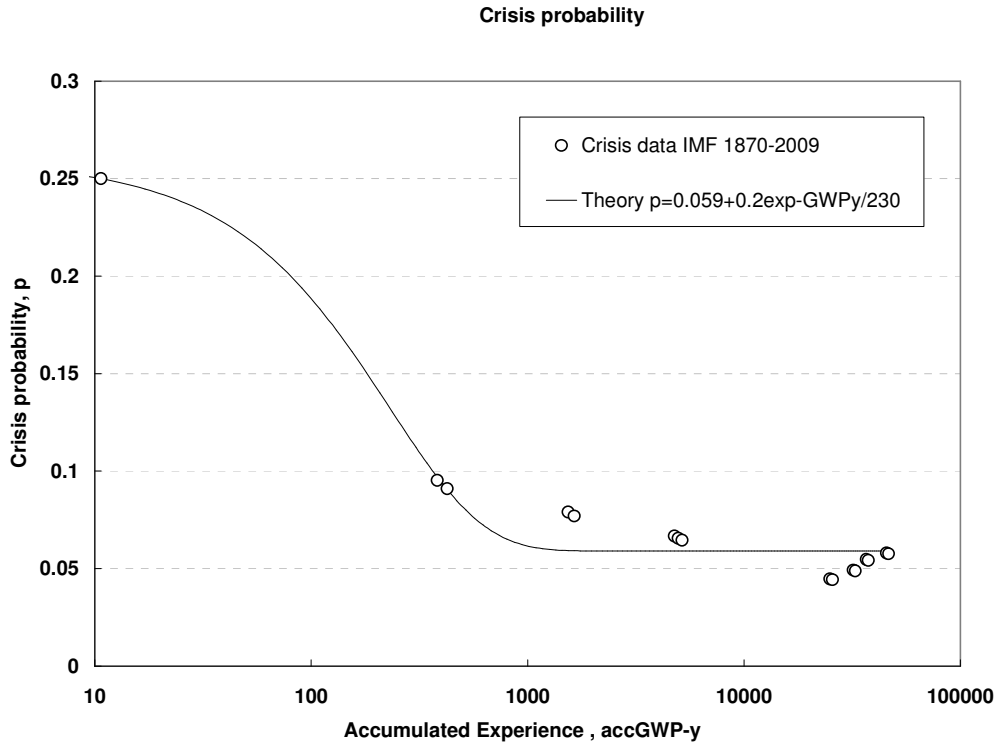
**Figure 8 The differential decline of rates of GDP growth**

What is seen is illuminating: the two growth rates (top dashed lines) are in anti-phase or negatively correlated: when one goes up, the other goes down, and vice versa. One grows literally at the expense of the other. There is also some periodicity in the divergence pattern, and evidence of emerging positive divergence in growth rates towards \$600T in 2003+. The opposite correlation between the growth rates is clear – the developing economies have a positive correlation of  $\sim +0.9$  and the developed economies a negative correlation of about  $-0.7$  with increasing accumulated GWP. As world wealth increases, one is declining, and the other is increasing in growth rate. The implication is that the relative growth shares part of the global economy “pie” growth, and hence the economies themselves are indeed closely coupled, which is perhaps obvious in hindsight.

The prediction is clear based on these trends. The developed world economies would actually go into near zero or into negative GDP growth rate in 2003+ (the projection is around 2005-2006 when GWP exceeds \$600T), after many years of decline. The emerging economies would continue to grow positively at 5% or more. The difference in rates was highly oscillatory and is perhaps not stable, as the liquidity (credit) needed to fund growth in emerging economies cannot come from those developed economies whose available assets and economies are in decline. So the implication is that – in a globalized economy where all the individual economies are linked – there are unknown feedback and stability relationships at work that we need to examine.

A very first attempt was also made to predict the actual probability of the known global fiscal crises, where the key is again finding the relevant units for the measure of the risk

exposure/experience,  $\tau$ . For the preliminary results shown in Figure 9, as listed in the IMF's WEO2009, the experience was taken as GWP-years for the interval 1870-2009 with eight non-wartime crises and the crisis probability as,  $p=(\text{Number of crises, per accumulated risk exposure years from 1870, accY})$ .



**Figure 9 Crisis probability estimate**

The theory line is derived from a MERE fit, so that the equation is:

$$p = 0.059 + 0.2 \exp(-\text{accGWP}/230),$$

with a correlation of  $r^2 = 0.958$ .

Clearly the predicted “tail” is constant at about 0.06 per year, suggesting a probability plateau from the finite minimum rate due to human involvement, as is evident from the comparison in Figure 9. Crises are occurring faster than might be have been expected using simple extrapolation with a power law. While not asserting any completeness at this stage of the analysis, it is possible in the future to further examine the trends in these data in more detail.

## System and Organizational Stability: SOS

How can we estimate the stability of a global or national system? The whole system is too complicated to predict its every move, behaviour or state: so how do we proceed? How can we estimate and predict the stability of a system when it is unpredictable? Here we introduce the only known objective measure of uncertainty, complexity and randomness, and illustrate how it can be used to predict system stability.

Markets and the entire physical world are random, chaotic and unpredictable, so predicting frequent and rare events is risky and uncertain<sup>5</sup>. Learning and randomness are powerful and unpredictable issues for risk prediction because we tend to believe that things behave according to what we know and, consciously or unconsciously, dismiss the risk what we have not seen or do not know about. After all, we do not know what we do not know. We, as humans, are the very product of our norms and patterns, our knowledge skills and experience, our learning patterns and neural connections, our social milieu and moral teachings, in the jobs, friends, lovers, lives, teachers, family and managers we happen to have. We perceive our own risk based on what we think we know, rightly or wrongly, and what *we* have experienced. But in key innovations and new disciplines, where knowledge and skill is still emerging – areas like terrorism, bioengineering, neuroscience, medicine, economics, computing, automation, genetics, law, space exploration, and nuclear reactor safety – we have to know and to learn the risk of what we know about what we do not know. We cannot possibly know everything, and these are all complex systems, with new and complex problems and lots of complexity, with much uncertainty.

The way to treat randomness and uncertainty has been solved in the physical sciences. Events can happen or appear in many different ways, which is literally the “noise” that surrounds and confuses us, whereas what we actually observe is the most likely but also contains information about the “signal” that emerges or is embedded, as order emerges from disorder, and we process and discard the complexity. In fact the whole process of individual human response time and decision-making has been shown to be directly affected by randomness, in the so-called Hick-Hyman law [2]. The number of different ways something can appear, or be ordered, in sequence, magnitude, position and experience, is mathematically derivable and is a measure of the degree of order in any system [2]. The number of different ways is a measure of the *complexity*, and is determined by the Information Entropy,  $H$ , which is also a measure of what we know about what we do not know, or the “missing information” [29], which is a measure of the risk. The relation linking the probability of any outcome to the entropy is well known from both Statistical Physics and Information Theory, and is the objective measure of complexity:

Information Entropy,  $H = \text{Sum} (p \times \text{natural logarithm, } p) = - \sum p \ln p$

---

<sup>5</sup>The inherent randomness is often termed the aleatory uncertainty by statisticians.



Note that the units adopted or utilized for the entropy are flexible and arbitrary, both by convention and in practice as being a comparative measure of order and complexity. So this measure of uncertainty requires a statement of probability. As individuals and as collectives, we do and must process complexity, both in our brains and in our behaviour, seeking the signal from all the noise, the learning patterns from the mistakes, the information from all the distractions. Now Taleb [1] noted in his notes that:

*“I am purposely avoiding the notion of entropy because the way it is conventionally phrased makes it ill-adapted to the type of randomness we experience in real life”.*

To make the entropy concept adaptable and useful for “experience in real life” all we have to do is actually relate and adapt the information entropy measure to our “real life experience”, or risk exposure interval, as we have already utilized [2, 3] and also introduced above. So we can now change the phrasing and the adaptability, since above we unconventionally phrased entropy as being “an objective measure of what we know about what we do not know, which is the risk”. In support of this new use and phraseology, other major contributors have remarked:

*“Entropy is defined as the amount of information about a system that is still unknown after one has made....measurements on the system” [30].*

*“This suggests that....entropy might have an important place in guiding the strategy of a business man or stock market investor” [11].*

*“Entropy is a measure of the uncertainty and the uncertainty, or entropy, is taken as the measure of the amount of information conveyed by a message from a source” [31].*

*“The uncertainty function...a unique measure for the predictability (uncertainty) of a random event which also can be used to compare different kinds of random events” [32].*

There is no other measure available or known with these fundamental properties and potential, particularly for handling uncertainty and randomness, the processing and influence of complexity, and providing the objective measure of order. This also has direct application to the subjective concept of ‘resilience engineering’, where ‘resilience is the intrinsic ability of an organisation (system) to maintain or regain a dynamically stable state, which allows it to continue operation after a major mishap and/or the presence of a continuous stress’ [33]. But ‘resilience’, just like “culture”, has not been actually measured or quantified anywhere: it is simply a desirable property. We have developed the numerical and objective system organizational stability (SOS) criterion that incidentally unifies the general theory and practice of managing risk through learning [2]. This criterion is also relevant to ‘crisis management’ policies and procedures, and emergency response centres in major corporations, facilities and industries.

The function of any “management system” is to create order from disorder, be it safety, regulatory, organizational or financial and hence to reduce the entropy. Hence, for order to emerge from chaos, and for stability in physical systems, the incremental change in entropy, which is the measure of the disorder, must be negative [34]. Our equivalent stability of organizational systems (SOS) criterion then arises imply from the fact that the incremental change in risk (information entropy, H) with changes in probability must be

negative, or decreasing with increasing risk exposure. In any *experience* increment we must have the inequality, expressed in differential form:

$$dH/d\tau \leq 0$$

This key condition requires that a maximum ('peak') exists in our changing missing information or state of order as a function of experience and/or risk exposure. To illustrate this variation, consider the limit cases of concern of the probability/possibility/likelihood of another collapse event, having observed a similar one already, considering all our previous knowledge. From the past experience we showed that the prior probability for repeat events (REP) is, with little learning,  $p \approx 1 - 1/e = 0.63$ , and also this value for novice mistakes with little experience (when  $\tau \rightarrow 0$ ). For the future risk, the posterior probability, with little learning, is  $p \approx 1/\tau$ , for rare events and also for highly experienced systems (when  $\tau \rightarrow \infty$ ).

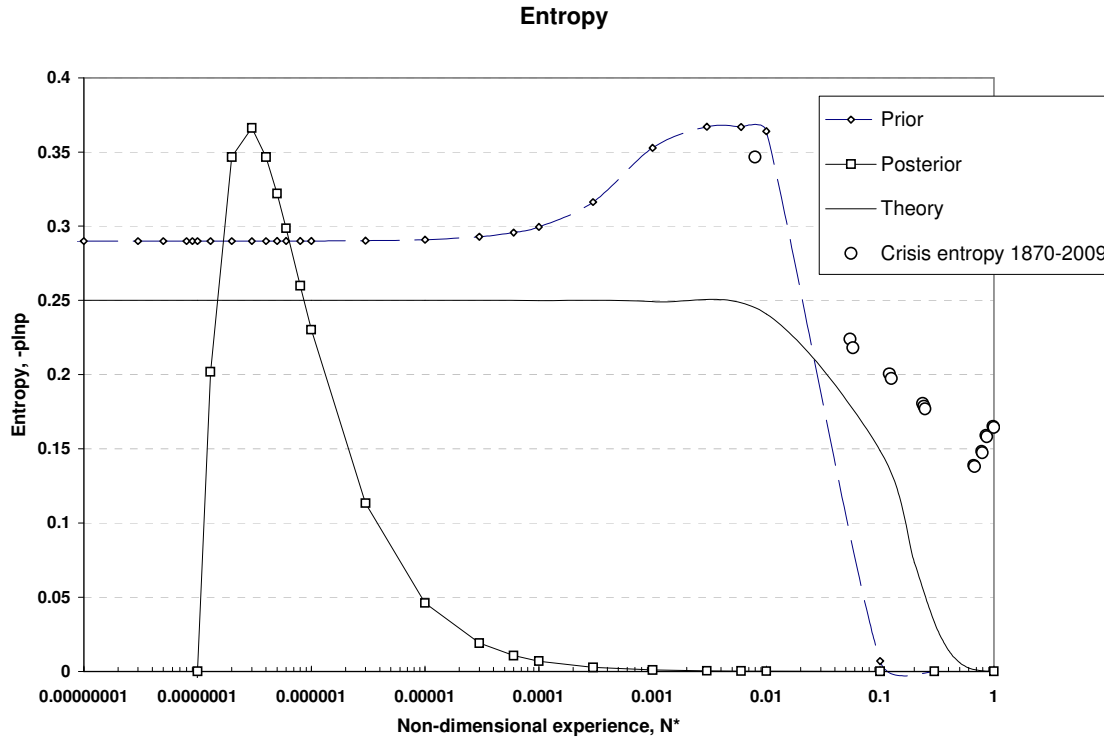
For the two *limited learning* cases of the prior (past MERE) and posterior (future rare event) the entropy increment,  $dH = -\ln p$  in any risk interval can be calculated. The results are shown in Figure 10 as a function of the experience or risk exposure interval,  $N^*$ , which purely for convenience has been non-dimensionalized to the maximum experience. For the example known "prior" case, entropy is calculated from the MERE probability result with little learning ( $k = 0.0001$ ); the decrease in entropy at larger experience or risk exposure in the prior case is due to the probability of an outcome finally reaching a certainty,  $p \sim 1$ , as ultimately there is no uncertainty. For the unknown "posterior" case, the entropy is calculated from  $p = 1/\tau$ ; the peak in entropy at small experience is simply due to the greater uncertainty, which decreases as experience is gained.

Also shown in the Figure 10 is the purely theoretical prediction obtained from SEST, the statistical error state theory [2], which treats outcomes as appearing randomly. The theory derives the most likely statistical distribution of outcomes, and relates the probability of the outcomes with variation in the *depth* of experience or risk exposure in any given risk interval. The information entropy,  $H$ , is the measure of the *complexity* in any interval and is given by integrating the resulting exponential probability distributions, to obtain:

$$H = \frac{1}{2} (p_0 \exp -aN^*)^2 (aN^* + \frac{1}{2})$$

At small experience, as  $N^* \rightarrow 0$ , the above SEST result becomes,  $H \rightarrow 0.25$ , which is close to the prior value with little learning of  $H \rightarrow 0.29$ , so the results are also consistent in their own limits as they should be. The value of the slope parameter or learning exponent,  $a$ , is derived deliberately from very diverse prior data sets for failure distributions which are very detailed and complete. Specifically, we used: (a) US commercial aircraft near mid-air collisions (NMACs) for 1987-1998, where experience is measured by total flights; and (b) Australian traffic fatalities from 1980-1999 where experience is measured in driver-years (as shown in [2, Figure 8.8]). The theory line in Figure 10 utilizes  $a=3.5$  in the above exponential distribution as a working approximate estimate for comparison purposes, which is close to the learning rate constant value,  $k \sim 3$ .

For most of the experience or risk range shown the entropy is not decreasing significantly until sufficient experience is attained.



**Figure 9 Entropy variation with experience, knowledge and risk exposure**

Knowns (prior or past) apparently contain more uncertainty ( $H$  is larger) than unknowns (posterior or future), except at very early or little experience ( $N^* < 10^{-4}$ ). The shapes are of interest for another reason, for evaluating the system organizational stability (SOS) criterion. By inspection of the two cases in Figure 10, this stability condition is only met or satisfied at small experience for the unknowns, and at large experience for the knowns.

Basically, at small experience *unless learning is occurring* the existing system is unstable and in danger of repeat events until very large experience is attained. Conversely, any future system is also initially unstable until sufficient post entry experience has been attained. So learning – or decreasing complexity – is *essential* for stability, and this is plainly relevant to the market stability when introducing the use of new and/or complex financial instruments.

The data points shown as circles in Figure 10 are for the trial “crisis entropy” estimates calculated using the prior preliminary probability values in Figure 9, where the risk interval or experience has been non-dimensionalized on the basis of the accumulated GWP from 1870-2009. The general data trend is downward (i.e. stable) until the last few data points for the crises of 1997 and 2007, indicating the potential for systemic

instability. This comparison suggests that entropy is indeed a potentially significant indicator that should not be simply avoided.

In our previous published work [2], we had quantified the uncertainty or complexity using the information entropy,  $H$ , as an objective measure of other subjective organizational and management desiderata of “safety culture” and “organizational learning” as a function of experience. This is the first time, to our knowledge, that entropy has been introduced as an objective prediction of the subjective feeling of “risk exposure” in the presence or absence of learning. Further work in this whole arena of system stability, the selection of relevant experience measures, and the quantification of future risk are clearly needed.

### **Concluding remarks**

An exercise such as predicting the “next” recession becomes simply equivalent to determining the probability of and risk interval for the “next” event or outcome. This probability must be based on relevant and correlated measures for experience and risk exposure, which include the presence or absence of learning.

We have summarised some recent ideas on risk *prediction* for multiple technological systems, using the existing data, and have explicitly included the key impact of human involvement using the learning hypothesis, namely that we learn from our mistakes. We have related these ideas to the prediction of rare events, systematic risk, and organizational stability in global systems and, although we do not pretend to have all the answers, there are clear directions to follow. Risk is caused by our uncertainty, and the measure of uncertainty is probability. The risk of an outcome (accident, event, error or failure) is *never* zero, and the possibility of an outcome *always* exists, with a chance given by the future (posterior) probability. The key is to include the human involvement, and to create and use the correct and relevant measures for experience, learning, complexity and risk exposure.

In industrial operations, the cardinal rule of operation applicable to *any* system is due to Howlett [35] which is:

*“Humans must remain in control of their machinery at all times. Any time the machine operates without the knowledge, understanding and assent of its human controllers, the machine is out of control.”*

Further the limits to operation are defined by a Safe Operating Envelope, with limits that include margins and uncertainty that define “guarantees” for the avoidance of failure. Risk management is then employed to protect or *mitigate* the consequences of failures that might occur anyway [35]. These well-tried concepts are all translatable to and usable in financial systems, just as they are for industrial systems, since *all systems include human involvement and hence involve the uncertainty due to risk taking and learning.*

The definition, use and concepts of risk adopted in the present paper utilizes measures for risk exposure and for uncertainty that encompass and are consistent with that proposed before [36]:

*“Risk entails two essential components:*

- *Exposure, and*
- *Uncertainty*

*Risk, then, is exposure to a proposition of which one is uncertain.”*

Standard statistical distributions and indicators presently used for financial systems (e.g., as used in VaR, or yield spread) are known to not be applicable to predicting rare events, systemic risk, crises and failures. Because of the human involvement, the risk becomes greater than just by using a Gaussian, normal or simple power law, until we reach very, very large experience and would have had a prior event anyway. We have a *greater* chance of outcomes and unexpected unknown unknowns if we are not learning than we might expect even from using simple “scaling” or “power laws”. *This is simply because we are humans who make mistakes, take risks and cannot be error free.* In colloquial terms, the human adds another “fat tail” to an already “fat tail”, which we may call a White Elephant.

So the past or prior knowledge indeed informs the future risk: what we know from what we already know from the probability of what were once past unknowns tells us about the probability of the unknown unknowns in the future, too.

The measure adopted and used and relevant for estimating risk exposure is key. Over some seven to eight decades (orders of magnitude) variation in the rate and in the risk exposure or accumulated experience, for the rare event the negligible learning prediction holds. At any future experience or risk exposure, the error in the prediction is evidently about a factor of 10 in future occurrence probability and about a factor of two in average frequency.

We have suggested several major factors and useful measures that influence the prediction of risk and stability in financial systems, based on what we observe for all other systems with human involvement:

- a) the Universal Learning Curve provides a comparative indication of trends;
- b) the probability of failure/loss is a function of experience;
- c) the relevant measure of success globally was the GDP growth rates;
- d) a relevant measure of experience is the accumulated GWP;
- e) stable systems are learning systems that reduce complexity;
- f) an absolute measure of risk and uncertainty is the Information Entropy;
- g) unique condition exists for systemic stability;
- h) repeat events are likely;
- i) existing systems are unstable unless learning is occurring; and
- j) new systems are unstable at small experience.

The rare events are essentially all the same, whether they be aircraft crashes, space shuttle losses, massive explosions, or huge financial crises: we know nothing about them until they actually happen, when and if they occur almost magically becoming “known unknowns”. We learn from them only after they have happened at least once. But based on what we know about what we do not know, we can always estimate our risk and whether we are learning or not. The rare unknown unknowns, or colloquially the “fat tails” or “black swans” of the unpredictable rate distributions, are simple manifestations of the occurrence of these outcomes whenever and wherever they happen. We can and must expect them to continue to appear, as Black Swans turn into White Elephants, time and time again.

The work and concepts discussed in this paper are only a first step in developing understanding for predicting and managing risk in complex systems with human involvement. This new application to financial systems and markets, and the adoption of new measures requires time, patience and can also introduce risk. Further work is clearly needed in this whole arena of system stability, the selection of relevant experience measures, and the quantification and prediction of future risk.

## References

- [1] Taleb, N., 2007, “The Black Swan: The impact of the highly improbable”, Penguin Books, UK.
- [2] Duffey, R.B. and Saull, J.W., 2008, “Managing Risk: The Human Element”, John Wiley & Sons Ltd., West Sussex, UK.
- [3] Duffey, R.B. and Saull, J.W., 2002, “Know the Risk”, First Edition, Butterworth and Heinemann, Boston, USA.
- [4] International Monetary Fund (IMF), 2009, “World Economic Outlook: Sustaining the Recovery”, Figure 3.1, October.
- [5] Columbia Accident Investigation Board (CAIB), 2003, “Report, Volume 1”, August, and “Working Scenario”, CAIB/NASA/NAIT, July (available at <http://www.caib.us>)
- [6] US Nuclear Regulatory Commission (NRC), 2008, “Davis-Besse Reactor Pressure Vessel Head Degradation Overview, Lessons Learned, and NRC Actions Based on Lessons Learned”, Report NUREG/BR-0353, Rev. 1, August (available at [www.nrc.gov](http://www.nrc.gov))
- [7] U.S. Chemical Safety and Hazard Investigation Board (CSB), 2007, “Refinery Explosion and Fire”, BP, Texas City, March 23, 2005”, Investigation Report No. 2005-04-I-TX.
- [8] Barthelemy, 2001, Report of the General Inspectorate for the Environment, Affair No. IGE/01/034. “Accident on the 21st of September 2001 at a factory belonging to the Grande Paroisse Company in Toulouse, produced jointly with the Explosives Inspectorate and with help from the INERIS”, 24 October, Inspection Générale de l'Environnement, Paris, France.
- [9] Report of the Commission of Inquiry, October 2007, Gouvernement de Quebec, available at [http://www.cevc.gouv.qc.ca/UserFiles/File/Rapport/report\\_eng.pdf](http://www.cevc.gouv.qc.ca/UserFiles/File/Rapport/report_eng.pdf).
- [10] Jaynes, E.T., 2003, “Probability Theory: The Logic of Science”, First Edition, Cambridge University Press, Cambridge, U.K, Edited by G.L. Bretthorst.

- [11] Rev. Bayes, T., 1763, Th. Bayes Memoir communicated by R. Price, "An essay toward solving a problem in the doctrine of chances", Royal Society of London, Vol 53, pp. 376-398.
- [12] Sveshnikov, A.A., 1968, "Problems in Probability Theory, Mathematical Statistics and the Theory of Random Functions", Dover, New York, pp.48, 49 and 80.
- [13] Rumsfeld, Donald, 2003, US Defense Secretary, quoted from press conference.
- [14] World Bank, 2009, available at [www.-wds.worldbank.org/WBSITE](http://www.-wds.worldbank.org/WBSITE), GWP 1981-2004.
- [15] Petroski, H., 1985, "To Engineer is Human: The Role of Failure in Successful Design", St. Martin's Press, New York, USA, Preface, p.xii.
- [16] Ohlsson, S, 1996, "Learning From Performance Errors", Psychological Review, 103, No 2, pp.241-262.
- [17] Coolen, F.P.A., 2006, "On probabilistic safety assessment in the case of zero failures", Proceedings, Institution of Mechanical Engineers, Part O, J. Risk and Reliability, Vol. 220, No. 2, pp. 105-114.
- [18] Duffey, R.B., 2004, "Innovation in Technology for the Least Product Price and Cost – A New Minimum Cost Relation for Reductions During Technological Learning", Int. J. Energy Technology and Policy, Vol. 2, Nos. 1/2.
- [19] Stevens, J.C. and Savin, H.B., 1962, "On the form of learning curves", J. Experimental Analysis of Behavior, 5,1, pp.15-18.
- [20] Hanayasu, S. and Sekine, K., 2004, "Damage Assessment of Industrial Accidents by Frequency-Magnitude Curve", Proceedings of PSAM 07, paper # 0229.
- [21] Woo, G., 1999, "The Mathematics of Natural Catastrophes", Imperial College Press, London, pp.224.
- [22] Duffey, R.B. and Ha, T.S. 2009. "Human Reliability, Experience and Error Probability: A New Benchmark", Proc. of the 17<sup>th</sup> International Conference on Nuclear Engineering", ICONE17-75861, 12-16 July 2009, Brussels, Belgium.
- [23] Fragola, J.R, 2009, "How safe must a potential crewed launcher be demonstrated to be before it is crewed?", Journal of Loss Prevention in the Process Industries, Vol 22, pp. 657-663.
- [24] Duffey, R.B. and Ha, T.S., 2009, "Predicting Electric Power System Restoration", Proc. IEEE (TIC-STH) Toronto International Conference – Science and Technology for Humanity, Toronto, Ontario, 27-29 September.
- [25] Duffey, R.B. and Tapping, R.L., 2009, "Predicting Steam Generator Tube Failures", Proc. 6<sup>th</sup> CNS International Steam Generator Conference, Toronto, Ontario, 8-11 November.
- [26] Duffey, R.B and Skjerve, A.B., 2008, "Risk Trends, Indicators and Learning Rates: A New Case Study of North Sea Oil and Gas", Proceedings ESREL 2008, 17th SRA Conference, Valencia, Spain.
- [27] International Monetary Fund (IMF), 2009, "World Economic Outlook of the International Monetary Fund", Update Table 1.1, available at [www.IMF.org/external/pubs/ft/weo/2009/update](http://www.IMF.org/external/pubs/ft/weo/2009/update), publication date Jan 28, 2009.
- [28] Estrella, A. and Mishkin, F.S., 1996, "The Yield Curve as a Predictor of U.S. Recessions", Current Issues in Economics and Finance, Vol 2, No 7, Federal Reserve Bank of New York, pp. 1-6.

- [29] Baierlein, R., 1971, "Atoms and Information Theory: An introduction to statistical mechanics", Dover, New York.
- [30] Wolfram, S., 2002, "A New Kind of Science", p. 44, Wolfram Media, Inc, Champlain, Ill., USA.
- [31] Pierce, J.R., 1980, "An Introduction to Information Theory", Dover, New York.
- [32] Grenier, W., Neise, L. & Stocker, H., 1997, "Thermodynamics and Statistical Mechanics", New York: Springer, pp. 150-151.
- [33] Hollnagel, E., Woods, D., & Leveson, N., (editors), 2006, "Resilience Engineering: Concepts and Precepts", Ashgate Publishing Limited, Hampshire, England.
- [34] Kondepudi, D. & Prigogine, I., 1998, "Modern Thermodynamics: From Heat Engines to Dissipative Structures", New York: John Wiley & Sons.
- [35] Howlett, H.C., 2001, "The Industrial Operator's Handbook", 2<sup>nd</sup> Edition, Techstar, Pocatello, Idaho, ISBN-10 1-57614-027X, pp. 40,63 and 76.
- [36] Holton, G.A., 2004, "Defining Risk", Financial Analysts Journal, Vol 60, 6, pp.19-25 ( [www.cfa.pubs](http://www.cfa.pubs) )