



# **Perfectly unpredictable: Why forecasting produces useful rubbish**

## **Professor Michael Mainelli**

### **28 January 2008**

Good evening Ladies and Gentlemen. I thought that it might be fun to be a little unpredictable myself, hence this Burns night hat!

As you know, it wouldn't be a Commerce lecture without a commercial. So I'm glad to announce that the next Commerce lecture will continue our theme of better choice next month. That talk is 'How To Get Ahead In Commerce: The Sure-Fire Ways To Make Money', here at Barnard's Inn Hall at 18:00 on Monday, 25 February.

An aside to Securities and Investment Institute, Association of Chartered Certified Accountants and other Continuing Professional Development attendees, please be sure to see Geoff or Dawn at the end of the lecture to record your CPD points or obtain a Certificate of Attendance from Gresham College.

Well, as we say in Commerce - 'To Business'.

#### Fracticality

At the heart of commerce is the notion that perfect markets aren't predictable, yet great efforts are expended on proving this notion wrong. 'Forecast' itself is an old word dating to at least the 14th century with the sense of scheming or contriving beforehand. Attempting to 'predict' or 'forecast' is not just human nature, but also a great industry. Forecasting is commonly used in supply chain management, weather, transportation planning, technology evaluation, experience curves, earthquakes, sports, telecommunications and demographics. Naturally, as all of these areas have a bearing on Commerce, forecasting is endemic in markets and economics. I can confidently forecast that people will be on television and in print tomorrow forecasting future market prices and economic conditions, though mostly to little avail. Perhaps we should all emulate Paul Gascoigne, 'I never predict anything, and I never will.' [The Economist 21 July 2007, page 17]

Why do we forecast? Interestingly, forecasting is normally about reducing future volatility. In systems theory, a basic system uses feed-back to correct itself and feed-forward (or positive feedback) to anticipate. The more accurate the feed-forward, the more efficient the system can be. In Commerce, if we could predict ever more precisely supply and demand, we could reduce volatility in the economy ever more as well. Better forecasting would reduce unnecessary shortages, lead times, over-stocking, under-investment, over-investment and waste in a myriad of ways. However, we encounter significant problems when forecasts are sporadically accurate. Patrick Young remarks, 'The trouble with weather forecasting is that it's right too often for us to ignore it and wrong too often for us to rely on it.' A more human view of

forecasting is that we have to:

'We will not perceive a signal from the outside world unless it is relevant to an option for the future that we have already worked out in our imaginations. The more 'memories of the future' we develop, the more open and receptive we will be to signals from the outside world.' [Arie De Geus, 1997]

[SLIDE: DIGITISING - CIRCA 1983]

In some regards, forecasting is simple, are things going to go up or down. Other kinds of forecasting, such as the shape of the future, can be much harder, or at least harder to describe. To get things started tonight, I'd like to start by looking at simply copying some lines from maps. Do pay attention because later on you'll have a chance to forecast some line directions yourselves.

Back in the early 1980's, I was directing the MundoCart and Geodat projects, the Google Earths of their day. We took maps from around the world, built software to digitise them and sold the data. Digitisation involved somebody tracing key features from the map. At the end of digitisation we had a series of points, excluding bathymetry or topography, something like 10,000 to 50,000 points per map. In those projects we were producing from six to twelve million points a year from hundreds of maps a year. It was hard, pernickety work and, if you made a mistake, you had to go back to the beginning of the line on which you were working. A 'typical' (no such thing in our experience) natural line, such as a coastline or a river bank, could contain a few hundred to a few thousand points, so you can appreciate the difficulty.

[SLIDE: NOT TO SCALE]

Having taken the points that made the lines, others would use our software to make maps at different scales. For example, one double-sided map covers the whole of the UK at 1:625,000, eight maps at 1:250,000, 204 maps at 1:50,000 and 403 maps at 1:25,000. One of our problems was rendering coastlines well. If one has a piece of digital coastline digitised at a scale of 1:50,000 and reduces it in order to display it at 1:625,000, it looks ugly, too complicated and wiggly. In addition it wastes computing power, storage and time. So people often remove points, 'generalising' it in the trade.

[SLIDE: DON't GENERALISE - FROM 491 TO 18]

While there were numerous experiments about the best way to do generalise, it rapidly became clear that just knocking out points didn't work well. In 1973, David Douglas and Thomas Peucker created a simple algorithm still used today that determines which points can be most easily removed without affecting the shape or 'feel' of the line. My sample here takes a line with 491 points, the line in black, and reduces it to just 18 points, the line in red - you can see all the discarded points to the right and left of the line. You can also see that the algorithm retains the character of the line rather well, despite reducing it's substance by some 96%.

[SLIDE: DIGITALITY]

The reverse of generalisation is to enlarge features, taking things digitised at 1:625,000 and displaying them at 1:50,000, where of course you wind up with very jagged and overly simplistic line. Engineers immediately begin to think that they can use spline functions to interpolate between the points, giving them nice smooth lines.

[SLIDE: ENGINEERS LIKE TO DO IT WITH CURVES]

Unfortunately, these smooth lines are just that, smooth lines. Why is smooth better than jagged? Smooth isn't more accurate. The extra points inserted to create the smooth line make it look as if the digitised line has been digitised at a much larger scale than was the case and any individual point is less likely to be accurate than a straight line between the known, digitised points.

[SLIDE: REAL VERSUS FAKE]

In the late 1970's and early 1980's fractals were all the rage, well all the rage if you were involved in computer graphics and computer cartography. Benoit Mandelbrot began his core book on fractal geometry by considering the question: 'How long is the coast of Britain?' It's a great question, because the answer very much depends on the length of the ruler you use. A yardstick gives a different answer than a foot and in turn a different answer than you'd get with a micrometer running around every grain of sand on the coast.

You can see here a coastline that was created artificially. Note the coastline's similarity to the coast of Australia at a low fractal dimension (0.11). A higher fractal dimension (0.36) makes things crinklier, more like the coast of Norway. Similar techniques are now commonly used for computer-generated imagery in two and three dimensions in many science fiction films and video games, as well as other applications. At the time we were desperate to use Mandelbrot's wonderful fractals for something, anything of practical use, and it turned out to be digital map enlargement.

[SLIDE: FRACTALITY]

What we did was to calculate the fractal dimension of the original coastline when it was digitised. When we needed to display a coastline digitised at 1:625,000 coastline at 1:50,000, we would simply insert a strip of fractal imagery of the same fractal dimension between the known points. People were cartographically amazed. If you zoomed in on Australia, the coast still felt like Australia. As you zoomed in on Norway, even up to absurd scales such as 1:6,250, i.e. 100 times larger than 1:625,000, it still felt like Norway, though you might only be looking at imagery generated around three or four known points. To a lot of people, this was wonderful, although a bit dangerous as you felt so certain that this really was the coast.

[SLIDE: SUPREMEY ENGINEERED FAKES]

To many engineers, this way of faking it, 'making it up', was somehow unethical. Things even got emotional. They felt that it was wrong to insert points in what seemed a non-interpolative, semi-random way just to make the line 'feel' better, yet they were happy to insert points through spline interpolation. When we pointed out that, on average, someone was just as likely to go out and find the real coast using fractal fake points as spline fake points, in fact in many cases more likely to find the fractal fakes were a better guess, they spluttered. Engineers preferred 'interpolation' accuracy to 'similarity' accuracy.

[SLIDE: ULSTERMATE ACCURACY]

There is a related problem in determining whether one map is more accurate than another. In the early 1980's I remember helping a woman with her post-graduate research into the accuracy of early modern maps of Northern Ireland. She wanted to see if different cartographic schools were markedly better or worse. It was a particularly humbling experience. Many of these old maps, though they seemed deformed

when overlaid on modern cartography, were quite accurate for certain purposes, particularly for distances from place to place. Again, these models of the world are difficult to assess on a single measure of accuracy. Accuracy could vary wildly across old maps, with some sections being highly detailed, and others close to declaring 'here be monsters'. But the maps were useful in their own time. As Pirsig noted about wider issues of quality, 'There isn't one way of measuring these entities that is more true than another; that which is generally adopted is only more convenient.' [Pirsig, 1974, page 237]

Now forecasting is related to coastlines. At the heart of forecasting is the concept of guessing which way lines are going to go in the future, but different people like to choose different methods. And it's difficult to assess which method is better.

## Forecasting Exercise

[SLIDE: LEFT OR RIGHT EXERCISE - 1]

As primates, one of our distinguishing characteristics among the animals is the large pattern-matching engine perched on top of our neck. Among vertebrates I hasten to point out that the unusual bit is the pattern-matching engine's size, not that it's on top of our neck. So we are going to have three little line exercises for both the pattern-matching engines on your necks, and your arms.

May I ask all of you to put both arms in the air. Can you leave the left arm up please and drop the right arm, or if that's not possible point to the left? Can you now drop the left arm and switch to the right arm up in the air, or if that's not possible point to the right? Great, you're now ready to play. The rules of this left/right game are simple. I'm going to reveal seven arrows. Each arrow points to the left or points to the right. Your exercise is to predict which way the arrow will point, right or left. Please keep your own score in your head as we proceed. I'm about to reveal the first arrow, so would everyone please choose whether the arrow will go right from this starting circle or left by putting a hand in the air?

This overly simple game doesn't count, it was just to get you going. Nevertheless, could I please have a show of hands from those who predicted all seven arrows correctly? [even those of you downstairs!] Six out of seven? We're going to play this exercise twice more. Before we do so, note that you probably have a basic algorithm in your heads - left followed by right, or right followed by left.

[SLIDE: LEFT OR RIGHT EXERCISE - 2]

Again, I'm going to reveal seven arrows. Each arrow points to the left or points to the right. Your exercise is to predict which way the arrow will point, right or left. Please keep your own score. I'm about to reveal the first arrow, so would everyone please choose whether the arrow will go right from this starting circle or left?

Now you realise that a simple left/right or right/left model isn't 100% accurate. But for some of you it was right six out of seven times. May I have a show of hands of those who predicted all seven arrows correctly? Six out of seven? A number of you probably thought I'd be rather nastier than I was and got caught up in your own machinations of 'what do I think he'd think I'd have thought he'd think I'd do', or somesuch.

[SLIDE: LEFT OR RIGHT EXERCISE - 3]

Again, I'm going to reveal seven arrows. Each arrow points to the left or points to the right. Your exercise is to predict which way the arrow will point, right or left. Please keep your own score. I'm about to reveal the first arrow, so would everyone please choose whether the arrow will go right from this starting circle or left?

May I have a show of hands of those who predicted all seven arrows correctly? Six out of seven? Five out of seven? Of course there were a few features that probably surprised some of you. You probably didn't expect the third arrow to go into the black. You probably didn't expect the fifth arrow to be reversed. You probably didn't expect to get the sixth arrow correct no matter what you did - assuming your hands were in the air. You probably didn't expect the last arrow to go straight down. We could get even more devious, for instance, not having the arrows link up, going up the screen and using many other devices that would demonstrate how complex we can make most games.

Now this left/right game is very simple. Some of you may even think I'm being a bit condescending. I'm not. Remember during this talk that you were emotionally involved in guessing, even if it was just to humour me. You were silently congratulating yourself on a correct guess and blaming me for appearing to change the rules in the third game. You were attempting to build a theory to predict the future.

[SLIDE: UNUSUAL SUCCESS]

A simple alternating left-followed-by-right or right-followed-by-left model (let's call this the Alt L/R model - 'alternating left/right' model) probably still got many of you a reasonable score. The Alt L/R model that worked in the first game perfectly and the second game almost perfectly, should have had little predictive capacity in the third game. Yet if you followed the Alt L/R model in the first game you would have had seven correct, in the second game five correct and in the third game also five correct. So the simple algorithm you might have used in the first game for 100% success would still have scored 71% on the second and third games. Alt L/R is even better than chance, with the addition of a simple rule that if the beginning line choice is wrong, repeat the same answer for the second line choice. Simple models of the world often work well, even when they don't really reflect the underlying reality.

But you probably feel, deep down, that while that Alt L/R may have performed equally well on the second and third game, somehow it shouldn't. In fact on the third game, the simple Alt L/R algorithm seems useless despite scoring well. It doesn't seem to fit a pattern, and your pattern-matching engine doesn't like this. There are of course a wealth of measures you can use to evaluate the success of an algorithm at predicting - mean absolute error, mean absolute percentage error, percent mean absolute deviation, mean squared error, root mean squared error, to name a few. These incur the challenge of determining which measure is valid or invalid in evaluating success, like the engineers and the interpolation or fractal approaches.

[SLIDE: SCIENTIFIC METHOD]

What distinguishes forecasting from pure guesswork is that you must have a hypothesis, otherwise it is just guessing. Being human, after the first game you had an Alt L/R hypothesis and were possibly about to accord it the status of a full scientific theory, on a par with relativity theory perhaps. Indeed, forecasting and prediction are central to the scientific method. One of my favourite quotes is from Samuel Karlin relates modelling to science, 'The purpose of models is not to fit the data but to sharpen the questions.' [11th R A Fisher Memorial Lecture, Royal Society, 20 April 1983] Though you shouldn't underestimate the power of guessing over knowing something in many circumstances:

'O. K. Moore tells of the use of caribou bones among the Labrador Indians. When food is short because of poor hunting, the Indians consult an oracle to determine the direction the hunt should take. The shoulder blade of a caribou is put over the hot coals of a fire; cracks in the bones caused by the heat are then interpreted as a map. The directions, indicated by this oracle are basically random. Moore points out that this is a highly efficacious method because if the Indians did not use a random number generator they would fall prey to their previous biases and tend to over-hunt certain areas. Furthermore, any regular pattern of the hunt would give the animals a chance to develop avoidance techniques. By randomising their hunting patterns the Indians' chances of reaching game are enhanced.' [Gimpl and Dakin, 1984, via Henry Mintzberg]

So, as well as requiring a hypothesis, true forecasting must incur the genuine risk of being wrong. You can't be a George Carlin weather forecaster, 'Weather forecast for tonight: dark.'. Always correct, but always useless. Still, too accurately, economic forecasting is often compared with divination of various forms.

[SLIDE: PHILLIPS' ECONOMIC COMPUTER]

I would point out to you that Commerce is full of scientific method issues, for example testing products, predicting sales, making investments or estimating inflation. And people work hard at forecasting. This slide shows Phillips Economic Computer in London's Science Museum:

'The 'plumber' responsible for this device was William Phillips. Educated as an engineer, he later converted to economics. His machine, first built in 1949, is meant to demonstrate the circular flow of income in an economy. It shows how income is siphoned off by taxes, savings and imports, and how demand is re-injected via exports, public spending and investment. At seven feet (2.1 metres) high, it is perhaps the most ingenious and best-loved of economists' big models.' [The Economist, 'Economic Models: Big Questions And Big Numbers', 13 July 2006]

How do you test a model's accuracy. You can look at every module within the model, but overall you're looking to see that the model 'shadows' reality taking into account the 'noise' in the measurements and the system. [Smith, 2007, page 111] But there are definitely problems assessing the accuracy of forecasts. What are we to do with Paul A Samuelson's observation that 'The stock market has forecast nine of the last five recessions.' Does a false positive forecast cost as much as a false negative? Or vice versa?

Financial Hurricanes

[SLIDE: FINANCIAL HURRICANES]

There is plenty of evidence that the quality of forecasting is poor. Sherden demonstrates that naïve algorithms, a bit like Alt L/R, often outperform sophisticated models. For example, he concludes that hurricane land-fall is better predicted using a ruler through the hurricane's two last known positions than using fancy meteorological models. But only rarely do people go back and actually evaluate old forecasts. This leads to the Billy Wilder remark 'Hindsight is always 20-20.' Or as Max Warburg pointed out in 1921:

'For most people suffer from exaggerated self-esteem, and specially bank managers, when they write their

annual reports three or six month later are inclined to adorn their actions with a degree of foresight which in reality never existed.'

[attributed to Max M Warburg in 1921, Joseph Wechsberg, *The Merchant Bankers*, Simon & Schuster, 1966, page 145]

Commerce loves to don the cloak of predict and control, analogous to the machine metaphors for organisations [Morgan, 1986]. Perhaps we ignore the evolutionary paradigm because it is so frightening:

'A picture emerges of strategies existing primarily to satisfy the psychological needs of the managers, in particular the need to feel in control. In reality, random mutations take place, and only the fittest survive. . . . The evolutionary perspective whilst intellectually appealing cannot be popular with business people. . . . Managers believe they ought to be able to accomplish something in the real world. The pure evolutionary paradigm has little or no forum in the organisational world.' [van der Heijden, 1996, page 35]

Strategic planning systems can be presumed to exist because people feel that they improve the future of, or for, the organisation. Our desire for strategic planning may be emotional; management emphasis is overwhelmingly directed at the future. The strategic plan is not judged on predicting the future, rather on feelings that our future has improved as a result of the process of planning. Firms that survive believe it's because they forecast successfully, yet it may be equally true that because they're the survivors their forecasts were, only with hindsight, accurate - worth thinking about. In the extreme,

'Is there really any difference between a chief executive asking his strategic planners for scenarios for the future and a Babylonian monarch making similar demands of his astrologer? We suspect not. Both may provide a basis for decisions where there is no rational method, given the current state of technology. But more important, both demonstrate that the leader is competent and knows what to do.' [Gimpl and Dakin, 1984]

Chaos and complexity feature frequently in talks about prediction and forecasting. 'There is nothing more certain about a complex system than to say it will be just like it is now a moment later.' [Kelly, 1994, page 435] At the same time, chaos and the science of complexity point to discontinuities that render the comfortable 'one similar moment following another' approach useless over reasonable time periods. The science of complexity does not ignore prediction, but examines prediction in different ways, e.g. as patterns to be extracted from appearances. Complexity theorists examine invariants (behaviours that recur despite radical system changes), growth curves (the ubiquity of applications for so-called S-curves) and cyclic waves (forms of repetition within complex systems). Nevertheless, the science of complexity holds little hope of improving predictability in the near term.

I don't want to recapitulate yet more absurdities of forecasting, rather I commend Sherden's book, *The Fortune Sellers: The Big Business Of Buying And Selling Predictions*, where he proceeds to demonstrate in most instances how poorly the \$200 billion prognostication industries of 1998 performed against their own standards in areas as diverse as weather forecasting, economics, financial analysis, demographics, technology forecasting, futurology and corporate planning. As John Kenneth Galbraith quipped, 'The only function of economic forecasting is to make astrology look respectable.'

So we conclude that we want to predict things, a lot, but it is jolly difficult to tell how good we are at it. Sometimes we generalise confidently from our experience, other times not. There are two major schools of statistics, frequentist, and Bayesian. The frequentist school is what most of us were exposed to at college, find a lot of examples of something and see how they are distributed. The Bayesian school says, in effect, have a good guess at what you think the distribution will be in advance and then modify it based on

experience. The Bayesian school can work with very little data. 'Priors' are the distributions people assume apply to the experiences they've had so far, even very little experience. When Hilaire Belloc wrote 'Statistics are the triumph of the quantitative method, and the quantitative method is the victory of sterility and death'. he clearly hadn't had the opportunity of exploring Bayesian statistics [BELLOC, Hilaire, 'The Silence of the Sea and Other Essays', 1940]

The normal or Gaussian distribution, 'the bell curve', is something that you use to estimate a lot of things, for example height, age and weight. If I said, 'what do you think is the weight of most 40 year old British men?', you would implicitly use a bell curve while if I said, 'what will the box-office takings be for the next Bond film?', you might use another. Griffiths and Tenenbaum published a fascinating pre-print in 2006 that described experiments aimed at identifying the priors people use for a variety of common sense questions. Their results suggest a 'close correspondence between people's implicit probabilistic models and the statistics of the world.' Striking a blow for common sense they even found that ordinary people guess telephone-queue waiting times using a Power Law distribution just as telecommunications engineers worldwide are slowly realising that this might be more accurate than their traditional Poisson distribution. Their and others research suggests that evolutionarily developed guessing mechanisms may be quite sophisticated, and adequate for many purposes. The Economist adds rather interestingly that this research:

'... might explain the emergence of superstitious behaviour, with an accidental correlation or two being misinterpreted by the brain as causal. A frequentist way of doing things would reduce the risk of that happening. But by the time the frequentist had enough data to draw a conclusion, he might already be dead'. [The Economist, 'Psychology: Bayes Rules', 5 January 2006]

Of course, The Economist should have pointed out that their conclusion about the likelihood of the frequentist being dead was itself Bayesian. The philosophical challenge remains - choosing between a successful inductive theory and a wrong one. Nelson Goodman posed the challenge back in 1954, 'The problem of induction is not a problem of demonstration but a problem of defining the difference between valid and invalid.' [Goodman 1954, page 77]. The experimental approach of science isn't always possible in Commerce. Once we admit that the trial is dangerous in some way, say jobs will be lost or customers will leave, then we are very unlikely to get people to agree to a scientific test. We could be more creative in many cases, for example conducting scientific trials of several ways to run schools or hospitals before we launch yet another new management approach, but it's never an easy sell. People want to select a solution, not test possible ones. This leads many organisations to follow what others do. A good example here is the assumption by many people in government and NGOs that the correct form of procurement is an open, competitive tender with all the bureaucracy and perverse results that seem to accompany public sector tenders. They rarely evaluate the alternatives used by many other organisations, and would find an experimental test terrifying, because their principal driver is conformity, not economy, efficiency, effectiveness, or even reducing corruption.

## Forecasting Forecasting

[SLIDE: FANS OF CHARTS]

I would like to move towards the conclusion of tonight's lecture with some observations on where forecasting may be headed - a forecast on forecasting. The rules for successful forecasters include, 'ideally, predict the past', 'make lots of predictions', 'predict a figure or a date, but never both', 'shout about your successes', 'leave town if your failures catch up with you'. A while back I pointed out that in a world of scaremongers and rational people, no one ever remembers rational people's predictions when they come

to pass, nor scaremongers' futures when they fail to arrive. People only remember scaremongers' successful predictions about the past. Today, technological advancement means that we are getting both better about the ways we forecast and about our memories, even if the latter is internet-powered (whatever the latter was).

My first observation is that we have already begun to present forecasts more sensibly. Probabilistic forecasts are more common and better understood. Instead of a single number that is almost certainly going to be wrong, with arguments in the future about how wrong it was, we see 'fan charts'. A good example is the Bank of England moving to fan charts for inflation in 1996. Expected outcomes spread out in bands from the most likely path of inflation. Actuaries are all excited about stochastic mortality models that provide fan charts. Fan charts are not new, dating back at least a couple of decades. A range of possible outcomes is no longer a sign of imprecision, it's a sign of maturity. John von Neumann would have supported fan charts, 'There's no sense in being precise when you don't even know what you're talking about.' Like the Bank, our firm, Z/Yen, has used fan charts for over a decade for sales forecasts. As many of you will know, I have long argued for fan chart, stochastic or confidence level presentations for many forward-looking measurements, in particular future company results.

Ranges of forecasted outcomes with their probabilities allow us to assess the accuracy of the forecaster over time. According to Richard Feynman, 'The first principle is that you must not fool yourself - and you are the easiest person to fool.' Presenting forecasts as probability ranges is a first step towards being honest with yourself and allowing you to evaluate your forecasting accuracy in the future. But it's not easy. Lenny Smith at the London School of Economics says, 'When we look at seasonal weather forecasts, using the best models in the world, the distribution from each model tends to cluster together, each in a different way. How are we supposed to provide decision support in this case, or a forecast?' [Smith, 2007, page 130] Lenny recommends more aggressive use of ensemble forecasts, iterations of a number of different models using states with different starting parameters. In our work [at Z/Yen], we've even suggested creating markets in the accuracy of forecasting models, for example political models used to predict the outbreak of civil war in nations.

[SLIDE: DYNAMIC ANOMALY AND PATTERN RESPONSE]

The second observation is often missed by those who judge computational advances from their desktop. The deep structure of computer systems is moving confidently towards dynamic anomaly & pattern response systems; these are dynamic - adaptive and learning from new data in real-time; anomaly - identify unusual behaviours; pattern - reinforce successful behaviours; response - initiate a real-time action; systems. New architectures support new applications for communities, trading, sales or customer relationship management. For example handling all of the incoming emails for an auction site or directing consumers to the goods they're most likely to buy, or appealing to a customer's likely charity or political affiliation. Increasing computational capability and increasing quantities of data give organisations the power to spot incipient trends early. Firms are taking advantage of their internet traffic of all forms, emails, website hits, orders and searches, to build increasingly fast, adaptive response systems. No successful global web owner can respond to commercial requests without automated systems. The volumes are too high.

In turn, Davenport and Harris [2007, pages 176-179] envisage more automated decisions, real-time analytics, visual presentation and machine-created strategies from transaction and text processing. Price isn't everything, particularly when you know so much about your customers and their likely responses. Of course, recommendations that jar can lead to hilarious consequences back at the consumer end. Writing in the Wall Street Journal in 2002, Jeffrey Zaslow's article, 'If TiVo Thinks You Are Gay, Here's How To Set It Straight', highlighted machines presuming too much, from their owners' sexual orientation in films to Amazon's public, embarrassing recommendation of pornography for its founder, Jeff Bezos. The emerging power of interacting directly with people's desires enriches, rather than undermines, our view of economics, though it is very likely to impair our privacy.

[SLIDE: PREDICTION MARKETS MATTER]

My third observation is that markets now matter more, and less. In the past, prices were the single most important piece of information. Prices about supply and demand were critical. Now, there are enormous prediction markets. The largest exchange by volume of transactions in the UK is not the London Stock Exchange, nor is the London Stock Exchange founded back in 1761 valued as highly as this exchange founded in 1999 - it is Betfair, processing bets between companies and individuals. H G Wells mused about the future, 'Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write.' I don't think that the past century has shown this forecast to be accurate, but intriguingly today's systems allow us to take the thinking of the citizenry and represent it statistically more easily than ever. You will know of many examples such as the University of Iowa's Electronic Markets for US elections set up in the 1990's, as well as markets such as the Hollywood Stock Exchange or the Foresight Exchange. These betting exchanges have moved from laboratory playthings, to being used inside companies such as HP for sales forecasts, to being fun public tools in their own right. Prediction markets supplement the real markets by anticipating movements. They are incipient trend detectors.

[SLIDE: QUALITATIVE VERSUS QUANTITATIVE]

Extrapolation or interpolation can be myopic, ignoring large-scale qualitative issues. Steve Davidson suggests that 'Forecasting future events is often like searching for a black cat in an unlit room, that may not even be there.' The fourth observation is that we increasingly have the ability to incorporate the qualitative with the quantitative. So-called 'imperfect-knowledge economics' strives to recognise the importance of qualitative regularities - patterns that are observable, durable and partially predictable. [The Economist, 'Economics Focus: A New Fashion In Modelling', 24 November 2007, page 108 - [http://www.economist.com/finance/displaystory.cfm?story\\_id=10172461](http://www.economist.com/finance/displaystory.cfm?story_id=10172461)] A good example might be the long-term forecasting of exchange rates based on purchasing power parity. The expectations of people that purchasing power parity matters affect long-term equilibrium, while the short-term equilibrium fails to materialise. Forecasters can use imperfect models, relying on large scale regularities in human nature and the environment to keep their model within bounds. Roman Frydman and Michael Goldberg have brought out a book on the subject, *Imperfect Knowledge Economics: Exchange Rates and Risk*, Princeton University Press, 2007.

In some ways we go back to the fractal versus the spline. If we start to lose coastline accuracy, it may be because we're moving from the coast of Australia to the coast of Norway. Quantitative models are great from day-to-day, but keep in mind the overall picture. Know when to throw them out and start afresh. Too often model failure induces a religious response, 'what we need to do is model harder!' When the quantitative models become less valid, we need to think more about using a fundamentally new model than necessarily fixing today's. Increasingly, thoughtful forecasters are using ensemble forecasts and scenarios to break out of day-to-day model loops. Instead of trying to find a best model, we try to build sets of robust models. Using scenarios one breaks out of today's models to build consistent pictures of alternative worlds based on different assumptions with new parameters. For example, in the London Accord work on climate change, the jump was to ignore day-to-day prices and models and just assume that the cost of CO2 would exceed €30/tonne. This new scenario suggested novel questions. Scenarios are more about recognising that prediction is futile, yet helping us identify the decision of most reward and least risk that provides the most future options.

[SLIDE: MODEL OR MODULE?]

The fifth observation paraphrases Marshall McLuhan, 'the model is the message'. Scientific and economic models have become the basis of speculative news. According to Victor Cohn, 'Scientists are to journalists what rats are to scientists.' To help you distinguish super models from scientific modules tonight, I

compiled this helpful slide. Yet more and more we find that scientific modules are the supermodels of the media. Michael Crichton's 2005 book, *State of Fear*, is an eco-thriller that has generated some controversy. Interestingly, the title refers not to the eco-scare plot, but to Crichton's underlying theme that 'politicized science is dangerous' and encourages the growth of scare-mongering in order to raise funds. Crichton points out that much of modern reporting is speculating. He contends that speculation is inexpensive media, quick quotes from talking heads. So long as unlucky speculations are forgotten and lucky speculations are remembered, this media output looks prescient. But he points to a deeper problem too. More and more media speculations are about model outputs, not real outputs. Controversially, Crichton chose climate change, using examples of news stories publishing the outputs of climate change prediction models.

However, this 'model is the message' thesis equally applies to interest rate models, inflation model, GDP models, credit models or any other model that is reported in the press. The model's outputs and future outputs can become the news, and a feed-forward loop has developed predicting with some accuracy what the 'true' news will be. A simple example is that some researchers use searches on the words 'depression' or 'recession' in the media to create leading indicators of business sentiment. Of course these researchers publish articles on the frequency of word occurrence that themselves increase the frequency of 'depression' or 'recession' in the media and influence people's opinions of what other people are thinking. Many indices and models, inflation, interest rate, financial centres' performance, GDP growth and others, affect people's perceptions. Widespread reporting on these models brings about self-fulfilling prophecies. At the beginning of this lecture we considered forecasting's principal benefit to be volatility reduction. Now we find that the very models we use increase instability. To paraphrase Goodhart's Law, 'when a model becomes a media model, it ceases to be a good model'.

[SLIDE: FORECASTING FORECASTING]

A paper by Ming Hsu et al [2005] indicates that ambiguity lowers the anticipated reward of decisions. In short, the more ambiguity about the potential odds in a choice, the more risk averse we may become. Ambiguity aversion is above and beyond being risk averse about long odds; we shun situations where we don't know the odds. So while you might gamble on a roulette wheel quite happily knowing the odds, you might find the ambiguity of betting on a wonky roulette wheel less appealing. This leads me to a closing thought.

Modelling is now more widespread than ever. Information technology makes models smaller, cheaper and faster, while the media make models news. As we saw tonight, models may not make you less stupid; they may make you stupid more abruptly. Forecasts from models have increased in importance. Rather than being simply analytical tools, they are social constructs. For example, imagine a businessperson sharing an investment decision with a friend, or a potential homebuyer talking at the office and you might overhear snippets such as - 'you're a fool to invest with this recession coming', or 'nobody ever lost money on bricks and mortar as long as these interest rates hold'. Forecasts have become core to societal group-think. Strangely though, the Caltech research suggests that strongly shared forecasts would reduce the perception of ambiguity, e.g., 'it's ok to invest now that I have confidence it is only going to be a shallow recession.' If this is the case, that widespread acceptance of forecasts is reducing ambiguity aversion while simultaneously the media model feed-forward effect on perceptions increases risk aversion, perhaps both effects are self-correcting.

Of course, true winners ignore all forecasts. 'Champions know that success is inevitable; that there is no such thing as failure, only feedback. They know that the best way to forecast the future is to create it.'  
[Michael J Gelb]

Thank you.

## Further Reading

1. CRICHTON, Michael, *State of Fear*, HarperCollins, 2005.
2. DAVENPORT, Thomas H and HARRIS, Jeanne G, *Competing on Analytics: The New Science of Winning*, Harvard Business School Press, 2007.
3. DE GEUS, Arie, *The Living Company: Growth Learning and Longevity in Business*, Nicholas Brealey Publishing, 1997.
4. DOUGLAS, David H and PEUCKER, Thomas K, 'Algorithms For The Reduction Of The Number Of Points Required To Represent A Digitized Line Or Its Caricature,' *Canadian Cartographer* Volume 10, Number 2 December 1973, pages 112-122.
5. GOODMAN, Nelson, *Fact, Fiction and Forecast*, Harvard University Press, 1954.
6. GIMPL, M L, and DAKIN, S R, 'Management and Magic,' *California Management Review*, Volume 27, Number 1, Fall 1984.
7. GRIFFITHS, Thomas L and TENENBAUM, Joshua B, 'Optimal Predictions in Everyday Cognition', preprint for *Psychological Science* 2006, <http://web.mit.edu/cocosci/Papers/prediction10.pdf> &<http://web.mit.edu/cocosci/josh.html> &[http://www.economist.com/displaystory.cfm?story\\_id=5354696](http://www.economist.com/displaystory.cfm?story_id=5354696)
8. MANDELBROT, Benoit B, *The Fractal Geometry of Nature*, W. H. Freeman and Company, New York, 1977 (1983 ed).
9. MILLS, Terence C, *Predicting the Unpredictable?*, Institute of Economic Affairs, Occasional Paper 87, 1992.
10. HSU M, BHATT M, ADOLPHS R, TRANEL D, CAMERER C F, 'Neural Systems Responding To Degrees Of Uncertainty In Human Decision-Making'. *Science*, 9 December 2005, Volume 310, pages 1680-1683 -<http://www.emotion.caltech.edu/papers/HsuBhatt2005Neural.pdf>
11. MORGAN, Gareth, *Images of Organization*, Sage Publications, 1986 (1997 ed).
12. ROSEWELL, Bridget, *Uses And Abuses Of Forecasting*, Futureskills Scotland, 15 March 2007 - [http://www.futureskillsscotland.org.uk/web/site/home/ExpertBriefings/Report\\_Futureskills\\_Scotland\\_Expert\\_Briefing\\_Uses\\_and\\_Abuses\\_of\\_Forecasting.asp](http://www.futureskillsscotland.org.uk/web/site/home/ExpertBriefings/Report_Futureskills_Scotland_Expert_Briefing_Uses_and_Abuses_of_Forecasting.asp)
13. SHERDEN, William A, *The Fortune Sellers*, John Wiley & Sons, 1998.
14. SMITH, Leonard, *Chaos: A Very Short Introduction*, Oxford University Press, 2007.
15. VAN DER HEIJDEN, Kees, *Scenarios: The Art of Strategic Conversation*, John Wiley & Sons, 1996.

## Further Surfing

1. Forecasting Principles - <http://www.forecastingprinciples.com/>
2. Playing with coastlines -<http://polymer.bu.edu/java/java/coastline/coastline.html>
3. Douglas-Peucker Algorithm -<http://www.simdesign.nl/Components/DouglasPeucker.html>
4. Henry Mintzberg on 'The Pitfalls of Strategic Planning' - [http://www1.ximb.ac.in/users/fac/dpdash/dpdash.nsf/pages/CP\\_Pitfalls](http://www1.ximb.ac.in/users/fac/dpdash/dpdash.nsf/pages/CP_Pitfalls)
5. The Economist, 'Bank of England: Always looking on the dark side?' 28 March 2002 - [http://www.economist.com/world/britain/displaystory.cfm?story\\_id=E1\\_TDVSPQJ](http://www.economist.com/world/britain/displaystory.cfm?story_id=E1_TDVSPQJ)
6. International Institute of Forecasters - <http://www.forecasters.org/>
7. Keith Kendrick, 'Why Do We Gamble And Take Needless Risks?', Gresham College, 2 March 2006 - <http://www.gresham.ac.uk/event.asp?PageId=108&EventId=371>
8. 'Prediction Markets: Does Money Matter?' - [http://www.newsfutures.com/pdf/Does\\_money\\_matter.pdf](http://www.newsfutures.com/pdf/Does_money_matter.pdf)
9. Phillips Economic Computer -[http://en.wikipedia.org/wiki/A.W.\\_Phillips](http://en.wikipedia.org/wiki/A.W._Phillips)
10. Jeffrey Zaslow, 'If TiVo Thinks You Are Gay, Here's How to Set It Straight', reprinted from the Wall Street Journal, 26 November 2002 -<http://www.mail-archive.com/eristocracy@merrymeet.com/msg00148.html>
11. University of Iowa Electronic Markets -<http://www.biz.uiowa.edu/iem/>

My thanks to Hamilton Hinds, Gerald Ashley, John Adams and Ian Harris for prompting ideas in this talk as well as, many years ago, T John Murray.

© Professor Michael Mainelli, Gresham College, 28 January 2008