The Boundaries of Quantitative Forecasting Methods: Respecting the Limits of Determinism

Adam Gordon

Abstract

Quantitative forecasting methods (QM) must overcome inherent challenges to be effective in guiding our future. These include modeling the four elements of relationship between inputs and outputs (shape, threshold, interactions, lags), recognizing the existence of chaos in the impacts of initial conditions, and dealing with wicked problems, mysteries, complex payoffs, and heavy tails. QM is valuable in situations of "Level 1 uncertainty," the prototype of which is a "mature industry" with a stable set of competitors, not under threat of technology upheavals or regulatory change, with unchallenged market segments and supply chains in a competitive environment that is well understood. QM is futile, however, for higher levels of uncertainty; rather than force application of these models into situations beyond their capability, we should concentrate on developing scenarios of alternative futures.

THE QM MYSTIQUE

Quantitative methodology (QM), including statistical analysis, has become the authoritative form of knowledge in modern Western societies. It stands at the core of the way we think about the world, how we account for what has happened, and how we anticipate what will happen next.

QM Is of Recent Origin This wasn't always so. QM is a relatively new phenomenon in the history of thought. It is only since the 19th century that numerical analysis has held this status, and alternative forms of investigation – judgment, experience, and intuition – have been pushed into the background. In our era, the pendulum has swung so far that, unless a view is numerically justified, it is almost not "knowledge." Economics, once an arena of social analysis, has become a field of turbo-math, while management academics produce papers that more closely resemble explications of particle physics than anything actual managers actually do. The pattern is repeated across much of psychology and the rest of the social sciences. The QM phenomenon has gained impetus with increasing access to computer modeling.

It is no surprise, then, that many people look to quantitative methods, and particularly computer-driven predictive modeling, to solve the conundrum of anticipating the future. All predictive models operate by processing algorithms that mathematically define relationships between variables, and set decision points about which variables affect which others, under what conditions, and by how much. With this in place, the modeler can project into future time at the touch of a button.

The Track Record Is Mixed The question, of course, is how good are these models at predicting? To what extent do predictive models actually prepare us for real future outcomes? The evidence is uncontested: quantitative models do well enough in situations where there are few surprises, and extremely poorly when the unexpected happens.

To improve the performance of predictive modeling, one may ask factual and bias-related questions about the base data, or technical questions about how algorithms are derived or assembled into a model, and seek ways to improve these procedures. There have been significant conceptual advances, including the "systems approach" to forecasting that, seeks to ameliorate many of the problems of traditional modeling by integrating agent-based models, network analysis, and system dynamics. See the feature section, "Rethinking the Ways We Forecast," in *Foresight*'s Summer 2009 issue.

But this is the terrain of much of the forecasting-methods literature. Our purpose here is to describe *the limits* of the validity of this mode of prediction; that is, where it is a valid enterprise and where it is not. Our question is not “How do we make a better predictive model for this situation?,” but “Is this the right situation for a predictive model?”

The problem, concisely, is predictive modeling very often bites off more complexity and uncertainty than any computer can chew. To see what predictive models are fundamentally up against, we have to investigate the problem of determinism and complexity in social systems.

DETERMINISM

Many systems in the world, particularly those found in physics or biology, are *deterministic*; that is, given certain starting conditions and certain forces, an outcome event is determined. By the laws of science, it will happen. If we heat water long enough, it boils.

The Scientific Method… The scientific method is about identifying, by experiment, which situations are deterministic, where the same inputs under the same conditions will always lead to the same outcomes (the experiment is “repeatable”). One may, for example, reliably predict the number of fruit flies in a cage tomorrow if the original number is known, because the number of factors that affect their fertility rate are reliably known, within a closed system, and will reliably apply every time.

…But Unscientific Human Behavior In the 18th century, social theorists thought the laws of human behavior were similarly deterministic – that a “science of society” was possible – and they set about uncovering these laws. It has since become clear that there aren’t deterministic laws governing human actions or society. It isn’t that we can’t see them, as classicists thought. They simply don’t exist. It is one thing to be able to determine the fertility of fruit flies, and altogether another to do this with regard to human and social situations that are multifactorial, where we do not and cannot understand all the variables influencing outcomes or the complex interaction between them.

In the science-of-society debate, philosopher Karl Popper definitively pronounced, "There can be no prediction of the course of human history by scientific or any other rational methods. … We must reject the possibility of social science that would correspond to theoretical physics" (~~Sherden, 1999,~~ Popper, 1957). Popper was defining a postclassical view of the world, one where we finally abandon the notion that, in human and social systems, a force or action will lead clearly to an expected effect or outcome, in every instance, repeatedly.

THE CHALLENGES OF PREDICTION

The Four Basic Elements In "Why We Can't Predict," ~~his~~ a short but appropriately severe essay, Neil Duncan (1992) divides the problems that must be solved to get to good algorithms – and therefore good forecasts – into four elements: *shape*, *thresholds*, *interactions*, and *lag*.

- *Shape* refers to the mathematical form of the relationship between input variables and outputs (the future predicted). The shape may vary from a simple straight line to a relationship that needs half a page of algebra to describe it.
- *Thresholds* are discontinuities in relationships, the points where the effect of an input factor suddenly changes. All values below a threshold may be zero, but above the threshold the factor influences events enormously and may cause a system to change entirely or collapse. (This same point was made by Roy Batchelor (2009) in a discussion of catastrophe-theory models, where a key variable falling below a critical value leads the whole system to respond in a way not predicted by a linear model.)
- *Interactions* are present when the effect of a factor depends on the values of one or more of the other factors. As "everything affects everything," our ability to forecast is commensurately reduced as the number of interconnections and interactions rises.

- *Lag* takes place when output is affected not by the current value of a factor but by an earlier or later value, raising the problem of how far back to go in considering pertinent inputs. (Lag itself may be subject to interactive or threshold effects.)

In any social system, each of these four categories of factors is often unknowable in itself, let alone in combination with others, making outcome possibilities unlimited.

Chaos Adding to the difficulty, chaos-theory analysts have postulated immense sensitivity to changes in initial conditions, the so-called *butterfly effect*, where change nuances or chance events can lead to widely differing outcomes when fully played out. A chance word or stray bullet can change history. As billions of small changes are happening around us all the time, it's hardly surprising that it is impossible to foresee outcomes to any reasonable standard of performance.

Wicked Problems Other ways of expressing irreducible complexity have also gained currency; for example, the concept of a "wicked problem," proposed by Rittel and Webber (1973) to describe situations that have incomplete, ill-defined, or contradictory interdependent variables. (A wicked problem contrasts with a *tame* problem where there is one clear solution, even if it's hard to find.) The point is there is *no* "solution" to a wicked problem.

Puzzles and Mysteries Along similar lines, Malcolm Gladwell (2007), citing U.S. national-security expert Greg Treverton, distinguishes between *puzzles* and *mysteries*. A puzzle is a problem that lacks information; if that information were provided, it could be solved. In mysteries, however, more information – or more information processing – doesn't necessarily help. As Gladwell says, "Sometimes the information we've been given is inadequate, and sometimes we aren't very smart about making sense of what we've been given, and sometimes the question itself cannot be answered."

Complex Payoffs and Heavy Tails In his essay "The Fourth Quadrant: A Map of the Limits of Statistics," Nassim Taleb (2008) weighs in on this debate by presenting a 2x2 matrix. On the one side are outcome (or "payoff") types, which are either "simple" (true/false) or "complex." On the other axis, outcome distributions are either "thin tailed" (outliers may occur, but carry little consequence) or "heavy tailed" (where outliers carry consequence and may overwhelm the data set).

Taleb argues that predicting via quantitative and statistical models will work adequately in the other three quadrants, but situations characterized by both complex payoff and heavy-tailed distribution are *beyond the limits of statistics*. We cannot apply quantitative projections to these situations because the existence of complex payoffs of potentially massive impact will make a laughingstock of the model. The situation is too complex and too uncertain to be predicted in this way.

The Grim Reality In sum, we seldom get adequate knowledge of initial conditions, nor a clear and exhaustive exposition of the forces driving the status quo, nor complete knowledge of how they interact. Therefore, there can be no science of human society or human institutions, and therefore no "scientific method" prediction of its future. Marketing analyst Pat LaPointe (2004) puts it in the most accessible terms: "Even though baseball statisticians have over 100 years of data loaded into high-speed computers at their fingertips, the human element in what happens with the very next pitch makes it nearly impossible to forecast (with any acceptable accuracy) who will win the game, never mind the pennant."

The analysts cited here, coming from philosophy, policy, finance, and beyond, are all saying essentially the same thing: complexity of human social systems limits the applicability of deterministic prediction models. Quantitative projections fail not because they are badly done, but because they are trying to do the impossible. It is more often than not beyond us to adequately capture the complexity of high-uncertainty situations *in the present*, let alone adequately forecast their evolution, and this is a problem cannot be "solved" by better algorithms, models, or computers.

FOUR LEVELS OF UNCERTAINTY

In a landmark essay in the Harvard Business Review, Hugh Courtney and colleagues (1997) set about asking: Is *uncertainty* an “all-or-nothing” phenomenon? They postulate four levels of uncertainty based on the amount of complexity and pace of change in the situation under study, as well as the time period under prediction – the further into the future we look, the more likely that key assumptions will shift over the forecast period. The levels are:

- Level 1, slow-moving, well-established situations where outcomes are dependable. The prototype of this is a “mature industry” with a stable set of competitors, not under threat of technology upheavals or regulatory change, with unchallenged market segments and supply chains in a competitive environment that is well understood.
- Level 2, where there is a limited, predetermined set of possible future outcomes, one of which will occur, but we can’t say which. This occurs in situations, for example, where there is an impending regulation that may go one way or the other, where a product may be approved or not (e.g., by the FDA), where there may be an industry-changing merger or not, or where different technology standards are fighting for adoption.
- Level 3, where outcomes are indeterminate but bounded within a plausible range that can be defined. Implausible or impossible outcomes can be identified and discarded. This type of uncertainty occurs in such situations as where new products or services are forming around new technologies or face uncertain customer demand, or where new delivery or revenue models are being tried. This type of uncertainty is also driven by changes in society, shifting values and norms, or unstable macroeconomic conditions (oil spikes, currency fluctuations, etc.).
- Level 4, where outcomes are unknown, unknowable, and there is a limitless range of possible outcomes. These are situations where, for example, new industries – e.g., genomics – are emerging, or where operating conditions are genuinely chaotic and unpredictable due to social or political factors.

Whether one parses uncertainty in this particular 4-level way or argues for fewer or more levels hardly matters. The insight is that uncertainty can be parsed at all. The further insight is that different uncertainty-management tools are appropriate for the different levels.

QM for Level 1 According to the authors, Level 1 situations are best analyzed with quantitative and statistical methods, including classic financial valuation (DCF or NPV). So, where we have low-uncertainty, slow-moving situations with a narrow scope and short time frames, it is perfectly reasonable and valid to forecast by algorithm. Put another way, the situation resembles a deterministic system well enough that quantitative modeling will return dependable results. In fact, this is the *best* way to think about the future – the QM analysis provides a disciplined and objective measure against which we can check our intuitions and experiential judgment.

Higher Levels For situations at higher uncertainty levels, more prone to reversals, changing assumptions, and heavy-tail outcomes, quantitative approaches quickly become inadequate. The moment a key assumption of the model fails, the model crashes. The problem is most situations in the modern world are medium- or high-uncertainty situations. To ask predictive modeling to make QM forecasts in these instances is effectively to disregard the real uncertainty level, inviting disaster. According to the authors, higher levels demand alternative tools (scenario analysis, game theory, decision trees, real options, etc.).

For example, let's say a property consortium wants to forecast hotel occupancy in the Toronto downtown area for the next five years. They start with quality data about existing occupancies over the past 50 years, along with a well-researched picture of tourist promotion business-conference initiatives and a thick wad of material from comparable areas of similar size and position. With these boosts, they should be able to come up with a quantitative future study that is accurate and helpful in making decisions. Sure, someone *could* decide to nuclear-bomb the city, rubbishing all predictions, but for practical, forward-planning purposes, it would be valid to say that assumptions of cause-and-effect between variables are known, dependable, and will remain good into the short-medium-term future – it is a Level 1 situation – so foresight culled through statistics and modeling is justifiable here.

However, if the same property consortium were to turn their attention to a 15-year future study of hotel occupancy in Havana, the quantitative process that served them well before would now serve very poorly. The situation is at a higher level of uncertainty. There may likely be political and other forces that open up this city to mass U.S. and worldwide tourism. It is quite possible to see Havana as the college "spring break" destination of choice in 2025, or to see it heavily influenced by Macau-based gaming mafia, or various other possibilities, all of which would make nonsense of carefully extrapolated models of present assumptions.

A simple example like this shows why, even as dramatic gains have been made in software and modeling, no quantitative forecast has ever proved adequate to the task of accurate prediction in high-uncertainty situations or in medium- or long-term time horizons, and none ever will. To attempt to determine the future in this way is no less than to disregard Popper and to perpetuate the 18th-century fallacy that society operates along scientific laws. This is not a problem that can be fixed by tweaking the model. Unfortunately, none of this seems to deter those who quantitatively predict, or at a minimum ~~even~~ inhibit them to Level 1 situations, not least because they can~~will~~ hide behind the promise of better software or more powerful computers that have~~should~~ "finally" cracked the problem. But,~~ no matter how sophisticated-seeming the quantitative approach, or how expensive the software, or how pretty the color images, ~~fundamentally unsolvable problems for QM prediction remain~~ underneath the swirl of projected numbers and their elegant presentation~~ ~~remain unsolvable problems for QM predictions~~ when asked to go beyond~~operating beyond~~ low-uncertainty situations, and, following Popper, they always will.

HOW TO PREDICT IN HIGH-UNCERTAINTY SITUATIONS

How, then, to think through high-uncertainty situations? ~~There is a way to proceed.~~ A beginning may be made by abandoning a modeling approach entirely for situations where such an approach is not valid. First, make a judgment of what level of uncertainty you are dealing with, and then use the analytical tools appropriate to that uncertainty. In the Havana example above, rather than futilely applying a tool designed for much more determined situations, and then planning around a spurious forecast, one should acknowledge that very different operating assumptions might come to pass, and explore how the future~~situations could evolve~~ may look under these different assumptions.

~~**Scenario Planning**~~ This is sometimes called "scenario planning," although this term has itself become a popular umbrella term for many forms of non-predictive foresight, and there is no attempt here to evangelize the scenario method. The point is, whether by scenario planning or by merely by exploring how the world would look under different operating assumptions, ~~without predicting noneany,~~ prior to modeling status quo operating assumptions, the foresight analyst starts by asking how good these assumptions are, and how likely they are to hold true into the future, and what new conditions may pertain. In other words, the analyst is precisely explore the assumptions that if (or more likely when) they were to change would destroy any quantitative model no matter how assiduously it was produced.

An assumption-testing mode of foresight does not produce number-exact forecasts of the type demanded by some managers. But it respects the residual uncertainty and capacity for fundamental change in uncertain and change-prone situations. Also, as it does not produce a spurious number, it does not induce the dangerous illusion of certitude when in fact there is none. As scenario planners say, "it is better to be vaguely right than exactly wrong."

Significantly, when using this framework – where the analyst acknowledges and respects that different operating assumptions may pertain – predictive modeling actually returns ~~again~~ to become a useful tool, even in very high-uncertainty situations. It allows data to be processed to flesh out the nature and implications of alternative, plausible model worlds. It adds ballast to the pictures of key alternative outcomes, allowing us to see a potential (but not predicted) plausible future more clearly, and therefore make better decisions going forward.

For example, at the time of this writing, various governments are involved in creating "stress tests" of banking-sector resilience. They model future situations, making different assumptions as to GNP growth, currency flows, borrowing rates, property prices, retirement age, etc., in an effort to foresee outcomes in each situation. This is quantitatively driven scenario planning, and the project is entirely valid because it makes no predictions. It merely uses the power of modeling to explore the outcomes implied in different operating or context assumptions. (Note that this is different from a "sensitivity analysis" where "high," "base," and low" forecasts are elaborated. Sensitivity analyses merely vary one key input metric, they do not vary fundamental assumptions.)

Earnest debates about the merits of different modeling methods exist, and each approach has its proponents and disciples. But seen from the point of view of appropriateness to uncertainty level, the difference between models is trivial. Subtly different quantification of the same unsound assumptions will not help. When forecasts of this type fail, it's never by some small margin that another model would solve. They miss by a mile, and would continue to miss if they used a competing model to project the future of the same wrong assumptions. As Steven Schnaars, commenting on the utterly failed projections for vertical-takeoff aircraft, icily observed, "A fancier version of the same model would have fallen deeper into the same hole" (1997, p. 158).

REFERENCES

R. Batchelor~~, R.~~ (2009), "Forecasting Sharp Changes" *Foresight,* Issue 13 (Spring 2009), pp. 7-12.

K. Popper (1957) "The Poverty of Historicism (Beacon Press, Boston) p158; cited in W. Sherden (1999), *The Fortune Sellers* (Wiley & Sons, New York), pp. 197-198.

N. Duncan; "Why can't we predict?" *New Scientist*, *136*(1841) (October 3, 1992), p 47.

H. Rittel, and M. Webber, "Dilemmas in a General Theory of Planning," *Policy Sciences, 4* (1973).

M. Gladwell, "Open Secrets," *The New Yorker* (January 8, 2007).

P. LaPointe, *5 Ways to Better Forecasts* (March 2004) Retrieved from www.marketingprofs.com, May 2008

N. Taleb, "The Fourth Quadrant: a Map of the Limits of Statistics," Edge Foundation - edge.org - September 15, 2008

[1] H. Courtney, J. Kirkland, P. Viguerie, "Strategy Under Uncertainty," Harvard Business Review, November 1, 1997

S. Schnaars, *Megamistakes* (New York: Free Press, 1989), p. 158.

BIO and PHOTO