

APRIL JOURNAL

CENTRE FOR FINANCIAL MANAGEMENT

CFM QUARTERLY IN FINANCE

APRIL 2022

EDITOR: DR. PRASANNA CHANDRA

CONTENTS

ARTICLE

How much do the experts know?

We have a tendency to construct and believe coherent narratives of the past. Everything seems to make sense in hindsight and this prods us to think that we can forecast the future. As Kahneman put it, "The illusion that we understand the past fosters overconfidence in our ability to predict the future."

The Illusion of Pundits

In his 2005 book, *Expert Political Judgment: How Good Is It? How Can We Know?*, psychologist Philip Tetlock examined the so-called expert predictions based on a landmark twenty-year study. He interviewed 284 people who made their living by commenting or advising on political and economic trends, both in their fields of specialisation as well as in the fields in which they had less knowledge. He asked the respondents to assess the probabilities that certain events would occur in the not-too-distant future. In every case, they were asked to rate the probabilities of three alternatives: the continuation of the status quo; more of something (economic growth or political freedom); or less of that thing.

The results were unequivocally bad: the experts performed worse than they would have, if they had simply considered each of the three potential outcomes as equiprobable. As Kahneman put it: "In other words, people who spend their time and earn their living, studying a particular topic, produce poorer predictions than dart-throwing monkeys who would have distributed their choices evenly over the options. Even in the region they knew best, experts were not significantly better than non-specialists."

It appears that those who know more predict very slightly better than those who know less, but those who know most are often less reliable. Why? A highly knowledgeable person develops an enhanced illusion of his or her skill and becomes unrealistically overconfident. As Tetlock put it, "We

reach the point of diminishing marginal predictive returns for knowledge disconcertingly quickly.” He added, “In this age of academic hyper-specialisation, there is no reason for supposing that contributors to top journals—distinguished political scientists, area study specialists, economists, and so on—are any better than journalists or attentive readers of The New York Times in ‘reading’ emerging situation.”

It is worth reiterating two lessons. One, since the world we live in is unpredictable, prediction errors are inevitable. Two, a high subjective confidence cannot be trusted as an indicator of accuracy – low confidence may be more informative.

Why Forecasting Thrives

If forecasts are really unreliable, why do people keep producing them? Perhaps experts don’t learn from their experience and ingeniously explain their failure. Financial analysts, for example, commonly say that they should not be evaluated on the basis of just one forecast (the single prediction defence) or something that happened outside of the scope of their model (the *ceteris paribus* defence).

Another reason is that it is a case of demand creating supply. When investment analysts and their managers are asked why do they engage in the futile exercise of issuing target prices, their last line of defence is always “Because the clients want them.”

We explored some reasons why people continue to produce forecasts, even if they are worthless. A bigger question seems to be: Why do people unthinkingly follow such useless forecasts? One view on this issue was articulated by Joe Nocera in an article that appeared in New York Times on October 1, 2005: “Indeed, I wound up thinking that forecasting is to the market what gravity is to the earth. As much as we like to poke fun at faulty predictions, we can’t function without them... Without forecasts, the market would no longer be grounded to anything.” I am not sure whether we need forecasts for investing. But Nocera suggests a reason why people keep using forecasts: When we are given a number we tend to cling to it, even without realising it. Cognitive psychologists call this phenomenon anchoring.

Superiority of Formulae, Models, or Algorithms

In a remarkably insightful book, *Clinical vs. Statistical Prediction: A Theoretical Analysis and a Review of Evidence*, psychologist Paul Meehl reviewed twenty studies that analysed whether clinical predictions based on subjective judgments of trained professionals were more accurate than statistical predictions based on combining a few scores or ratings according to a formula. He found that statistical predictions were better than clinical predictions. This book provoked considerable controversy and engendered a stream of research that is still continuing even after more than sixty

years of its initial publication. Nearly two hundred such studies done over decades have shown that algorithms are significantly more accurate than humans.

Similar studies have been done in other domains such as medicine, economics, finance, public policy, and sports. Each of these domains is characterised by a high degree of uncertainty and unpredictability. They are referred to as “low-validity environments.” Interestingly, in every case, the accuracy of experts has been matched or exceeded by a simple algorithm. As Meehl said with justifiable pride three decades after the publication of his book, “There is no controversy in social science, which shows such a large body of qualitatively diverse studies coming out so uniformly in the same direction as this one.”

Why are experts inferior to algorithms? One reason is that experts try to consider complex combinations of features. While complexity may work in an odd case, more often than not it reduces validity. Another reason is that humans are incorrigibly inconsistent in drawing inferences from complex information. In his revolutionary book, *The Limits of Scientific Reasoning*, David Faust wrote: “Human judgment is far more limited than we think. We have a surprisingly restricted capacity to manage or interpret complex information.” Studying a wide range of professionals, Faust found that simple quantitative models consistently outperformed human judges.

As James P.O. Shaughnessy put it in his book *What Works on Wall Street*: “Models beat human forecasters because they reliably and consistently apply the same criteria time after time. In almost every instance, it is the total reliability of application of the model that accounts for its superior performance. Models never vary. They are always consistent. They are never moody, never fight with their spouse, are never hung over from a night on the town, and never get bored. They don’t favour vivid, interesting stories over reams of statistical data. They never take anything personally. They don’t have ego. They’re not put out to prove anything. If they were people, they’d be the death of any party.”

Robustness of a Simple Algorithm

Since Meehl’s original work, the most important development in the field has been Robyn Dawes’ famous article, “The Robust Beauty of Improper Linear Models in Decision Making,” published in a 1979 issue of *American Psychologist*. The dominant practice in social sciences is to use an algorithm, called multiple regression analysis, for assigning weights to different predictors (factors). Multiple regression analysis has a sound logic and it determines the optimal formula that puts together a weighted combination of the predictors. However, Dawes observed that a formula in which predictors are equally weighted is likely to be just as reliable in predicting new cases as the multiple regression formula derived from the original sample. More research goes a step further: equal-weighted formulas are superior to multiple regression formulas derived from the original sample, because they are not affected by sampling errors. As Kahneman observed in his seminal work, *Thinking, Fast and Slow*, “The surprising success of equal-weighting schemes has an important practical implication: it is possible to develop useful algorithms without any prior statistical research.

Simple equally weighted formulas based on existing statistics or on common sense are often very good predictors of significant outcomes.”

In a memorable example, Dawes showed that marital stability can be predicted by a simple formula:

Frequency of lovemaking – Frequency of quarrels

The moral of the story is that a simple, back-of-the envelope algorithm is often as good as an optimally weighted formula and certainly better than expert judgment. This logic is applicable in many domains, ranging from stock selection to choice of medical treatment.

A classic application of this approach is a simple method for assessing the condition of a new-born baby. Developed by Dr. Apgar, this method says examine five variables (heart rate, respiration, reflex, muscle tone, and colour), assign each of them one of the three scores (0, 1, or 2, depending on the robustness of each sign), and rate the infants by this rule one minute after they are born. A baby with a total score of 8 or above is likely to be in good shape and a baby with a total score of 4 or below needs immediate intervention. The Apgar test has made an important contribution to lowering infant mortality. Atul Gawande’s book, *A Checklist Manifesto*, provides many other examples of the benefits of checklists and simple rules. As Atul Gawande said, “In domain after domain, aviation, medicine, construction, and investing where people have properly implemented a checklist, the outcomes have improved without improvement in the underlying skill of the users.”

The Hostility to Algorithms

Clinical psychologists reacted to Meehl’s ideas with hostility and disbelief. The statistical evidence of clinical inferiority is incongruent with clinicians’ experience of the quality of their judgments. Psychologists who work with patients have many hunches during therapy sessions that are confirmed, suggesting the reality of their clinical skill.

The problem is that the correct judgments relate to short-term predictions (a skill in which therapists may have years of practice). However, they typically fail at tasks that require long-term predictions about the future of patients.

It appears that the debate about the virtues of clinical and statistical prediction has always had a moral dimension. As Meehl wrote, experienced clinicians criticise the statistical method as “mechanical, atomistic, additive, cut and dried, artificial, unreal, arbitrary, incomplete, dead, pedantic, fractionated, trivial, forced, static, superficial, rigid, sterile, academic, pseudoscientific, and blind.” On the other hand, the proponents of the clinical method lauded it as “dynamic, global,

meaningful, holistic, subtle, sympathetic, configural, patterned, organised, rich, deep, genuine, sensitive, sophisticated, real, living, concrete, natural, true to life, and understanding.”

When Can You Trust Expert Intuition

Kahneman has been the leader of the heuristics and biases approach. This approach is not liked by Gary Klein, the intellectual leader of an association of scholars and practitioners who call themselves students of Naturalistic Decision Making or NDM. They work mostly in organisations where they study how experts work. Gary Klein wrote a book, titled *Sources of Power*, in which he analyses how experienced professionals hone intuitive skills. They are critical of the heuristics and biases approach because they think it is overly concerned with failures and based on artificial experiments rather than the study of real people doing things that really matter. They are skeptical about the use of algorithms in place of human judgement.

Interestingly, the leaders of the two groups, viz., Kahneman and Klein engaged in an “adversarial collaboration” to map the boundary that separates the marvels of intuition from its flaws. After several years of discussions and argumentation they published a joint article, titled “Conditions for Intuitive Expertise: A Failure to Disagree.” They concluded that their disagreement was partly due to the fact that they had different experts in mind. Klein spent much time with clinical nurses, fireground commanders, and other professionals who have real expertise. Kahneman had spent more time thinking about clinical psychologists, stock pickers, and political scientists trying to make unsupportable long-term predictions. Understandably, Klein has a lot of trust in expert intuition, whereas Kahneman has skepticism.

To understand when judgments reflect true expertise and when they display the illusion of validity, we should know the conditions for acquiring a skill. It seems that skill can be acquired when the environment is sufficiently regular to be predictable and there is an opportunity to learn these regularities through sustained practice. When both the conditions are fulfilled, intuitions are likely to be skilled. Chess is an excellent example of high validity environment where there is an opportunity to learn through sustained practice. In contrast, stock pickers operate in a zero-validity environment.

Expert Squeeze

As networks draw on the wisdom of crowds and computing power increases, the ability of experts to improve predictions is steadily diminishing. Michael Mauboussin calls this the expert squeeze and the evidence for the phenomenon is mounting.

Experts do have an edge in some crucial areas. Mauboussin provides guidance, as given in the box below, to know when and how to use experts.

Snippets

1. Mental Models

Psychologist Kenneth Gaik in his work *The Nature of Explanation* wrote that people construct mental models for processing information. He believed that with a “small- scale model of external reality and of possible actions” in our head, we “try out various alternative, conclude which is the best of them, react to future situations before they arise, utilize the knowledge of past events in dealing with the present, and in every way react in a much fuller, safer, and more competent manner to the emergencies which [we] face.”

Psychologist Kenneth Gaik in his work *The Nature of Explanation* wrote that people construct mental models for processing information. He believed that with a “small- scale model of external reality and of possible actions” in our head, we “try out various alternative, conclude which is the best of them, react to future situations before they arise, utilize the knowledge of past events in dealing with the present, and in every way react in a much fuller, safer, and more competent manner to the emergencies which [we] face.”

Through a series of controlled experiments, Princeton psychologist Philip N. Johnson- Laird examined how people construct mental models which he detailed in his book *Mental Models*. He observed that humans commit several systematic errors:

We tend to assume that each model is equiprobable. Humans do not seem to be mentally wired to perform Bayesian inferences

When people have a set of mental models about a particular phenomenon, they tend to focus on just a few- and sometimes only one. This can obviously lead to erroneous conclusions.

Mental models typically represent what is true but not what is false. For example, it is easier to construct a mental model of what inflation is rather than what it is not.

People have a distressing tendency to create mental models based on superstition and unwarranted beliefs.

What makes people so gullible? In his book *How We Believe* , psychologist Michael Shermer argues that the answer lies in the power of belief system. We seek patterns because our survival as a species has depended on this ability. As Shermer says, “Those who are best at finding patterns (standing upwind of game animals is bad for the crops) left behind the most offspring [and] we are

their descendants.” We are hardwired to seek patterns to explain our world and these patterns are the foundation of our belief system, even when they are specious.

The Age of Science diminished the errors in thinking and irrational beliefs. But we still succumb to magical thinking because as pattern seeking animals we need explanations even for the what cannot be explained.

As Robert Hagstorm says, “Our brains are belief engines that naturally look for patterns, which are then infused with meaning. Not surprisingly, we look for information that confirms our beliefs while ignoring information that contradicts them. Shermer refers to this “belief dependent reality.” Perhaps the age- old phrase “seeing is believing” has to be turned on its head, “believing is seeing.”

2.Narratives and Numbers

We communicate with each other basically through a series of stories. Stories are essentially open-ended and metaphorical and not determinate. Now scientists philosophers, doctors, journalist, and others have substituted the word “narrative” for the word ‘story –telling.”

Investors too have narratives. However, the narratives investors use to explain the economy or market sometime lack statistical rigour required for a proper description. And if the description is flawed the explanation is likely to be wrong.

Mathematician John Paulos tells us that people are good at storytelling (or narratives) and also decent at statistics. However, rarely does the storyteller invoke a statistical defence for the story. Likewise, while people cite good statistics, they rarely put statistics into a proper context. Says Paulos, “Unfortunately, people generally ignore the connections between the formal notions of statistics and the informal understanding and stories from which they grow.” He adds, “They consider numbers as coming from a different realm than narratives and not a distillations, complements, or summaries of them. People often cite statistics in bald form, without the supporting story and context needed to give them meaning.”

According to Paulos, when we listen to stories we tend to suspend belief in order to be entertained and when we evaluate statistics we are less willing to suspend belief, lest we are duped.

There are two types of errors in formal statistics.

Type I Error: We observe something that is not there.

Type II Error: We fail to observe something that is actually there.

Those who like to be entertained and wish to avoid making. Type II errors tend prefer stories over statistics. As Paulos explains, "The focus of stories is on the individual rather than the averages, on motive rather than movements, on context rather than raw data." Those who are keen to avoid Type I errors and have no great yearning for stories tend to prefer statistics to stories.

3. Best Strategic Thinkers

People commonly associate strategy with rational thinking and other high level functions of the prefrontal cortex such as anticipation, probability assessment, risk appraisal, pattern recognition, and abstract thinking. But researchers have found that best strategic thinkers show more activity in parts of the brain such as the insula, the anterior angular cortex, and the superior temporal sulcus, which are associated with emotion, intuition, empathy, and emotional intelligence.

Humour

1. Terms and Conditions

In an iron country a young peasant went to join the army. The recruiting officer asked "Which Regiment do you wish to join?" The peasant said, "In the General Staff."

The officer, "Are you mad?"

The peasant, "Why is that a pre- condition?"

2. Adam and Eve

A Russian teacher asked a student "Who were the first human beings?"

Student "Adam and Eve."

Teacher "What was their nationality?"

Student "Russian, of course."

Teacher "How do you know?"

Student, "It is so simple. They had no roof over their heads, no clothes over their body, and only apple for the two of them. Yet they called it Paradise."

Wise Saws

A great city is one that handles its garbage and art equally well.

Politics is like roller skating. You go partly where you want to go and partly where the dammed thing takes you.