Neapolis University

HEPHAESTUS Repository **http://hephaestus.nup.ac.cy**

School of Economic Sciences and Business Articles

2016

Forecasting and uncertainty: A survey

Makridakis, Spyros

IOS Press

http://hdl.handle.net/11728/7334

Downloaded from HEPHAESTUS Repository, Neapolis University institutional repository

Risk and Decision Analysis 6 (2015/2016) 37–64
DOI 10.3233/RDA-150114
IOS Press

Forecasting and uncertainty: A survey

Spyros Makridakis [a,*] and Nikolas Bakas [b]
[a] *Rector Neapolis University Pafos, Paphos, Cyprus*
[b] *Professor Neapolis University Pafos, Paphos, Cyprus*

Abstract. The origins of forecasting can be traced back to the beginning of human civilization with attempts to predict the weather, although forecasting as a field first appeared in the 1940s and attracted more followers from the early 1950s, when the need for predictions emerged in different fields of endeavor. It expanded considerably in the 1960s and 1970s when benefits were ascertained and computers were employed to perform the tedious calculations required. But initial successes in the fields of economics and business were first moderated and later reversed, with reality checks, first during the 1973/74 energy crisis, afterwards during the prolonged economic stagflation of the late 1970s and early 1980s and further deteriorated during the severe 2007/8 global financial crisis. The initial, optimistic expectations that social sciences will (using powerful computers and sophisticated models) replicate the predictive accuracy of hard ones were repeatedly shattered. This has left diverse fields like economics, management, political and human sciences and even worse medicine with no objective evidence of successful, accurate predictions, casting doubts to their usefulness and "scientific" vigor. At the same time, weather forecasting achieved success for immediate term predictions improving its accuracy and reliability over time. This paper starts with a historical overview of non-superstition based forecasting as it is practiced in different areas and surveys their predictive accuracy, highlighting their successes, identifying their many failures and explaining the reasons involved. Consequently, it argues for a new, pragmatic approach where the emphasis must shift from forecasting to assessing uncertainty, as realistically as possible, evaluating its implications to risk and exploring ways to prepare to face it. This paper expands Rumsfeld's classification to four quadrants (Known/Knowns, Unknown/Knowns, Known/Unknowns and Unknown/Unknowns) in order to explore the full range of predictions and associated uncertainties and be able to consider the implications and risk involved. Finally, there is a concluding section summarizing the findings, reiterating that we must accept that the great majority of our predictions will be inaccurate and that uncertainty, sometimes huge, surrounds all aspects of our future. There are also some suggestions for future research aimed at turning forecasting into an interdisciplinary field increasing its value and usefulness.

Keywords: Forecasting, uncertainty, risk, time series, econometrics, averaging, decision rules, judgmental forecasts, prediction markets, simple versus sophisticated models

> "To teach how to live with uncertainty, yet without being paralyzed by hesitation, is perhaps the chief thing that philosophy can do".
>
> Bertrand Russell

> "We really can't forecast all that well, and yet we pretend that we can, but we really can't".
> Alan Greenspan (whose belief in accurate predictions was partly responsible for the 2007/2008 Great Recession)

Over much of human history, the desire to predict the future was motivated by fear and was grounded on superstition and speculations of supernatural forces. Priests, augurs, oracles, seers and soothsayers, among others, fulfilled the role of forecaster, and were often rewarded handsomely for their services. Despite decades of scientific research on forecasting, little has changed in forecasting practice. Commonly used methods are rooted on folklore, unproven facts and even superstition. The human desire to know, with certainty, how things will turn out is unchanged, and drives a demand for forecasts that cannot be satisfied by evidence-based findings. Superstition-based forecasts, provided by "experts" with their guru reassurance of confidence verging on certainty are common and widely used.

We believe that the time has come to evaluate the accumulated evidence and to establish, in a realistic and objective way, what can and cannot be predicted and with what certainty. This paper is organized into four parts. First, we provide a short history of non-superstitious forecasting, and survey the post-WWII era including the efforts to consolidate fore-

*Corresponding author. E-mails: Spyros.makridakis@nup.ac.cy, Spyros.makridakis@gmail.com.

1569-7371/15/16/$35.00 © 2015/2016 – IOS Press and the authors. All rights reserved

casting into a single field. Second, we document successes and failures of the forecasting efforts in various disciplines during the modern era. Third, we argue the need for a new pragmatic shift in objective, from accurate forecasts to realistic assessment of uncertainty, in order to facilitate the evaluation of risks and be prepared to face the future rationally. We do so by expanding Rumsfeld's classification of predictions and uncertainty into four quadrants, namely: Known/Knowns, Unknown/Knowns, Known/Unknowns, and Unknown/Unknowns. Finally, we summarize the implications of our survey's findings on the limits of predictability and uncertainty, and propose the creation of an integrated, multidisciplinary forecasting field that can provide useful information for decision and policy makers.

1. The inability to forecast and the resulting uncertainty

On December 26, 2004 a massive tsunami, with waves up to 50 meters high, hit the south-eastern coasts of Asia causing 230,000 deaths and devastation over large areas. The previous known deadliest tsunami had hit Lisbon on November 1, 1755, with waves of up to 30 meters causing 60,000 deaths. There were no historical records that would lead one to expect the sudden disaster that was caused by the 2004 tsunami. A little more than six years later, another giant tsunami hit the north pacific coast of Japan flooding huge coastal land areas and causing a serious meltdown on the Fukushima Daiichi Nuclear Power Plant as the Plant was only built to withstand tsunamis of less than fifteen meters.

On September 15, 2008 Lehman Brothers, a big US bank with $639 billion in assets and $619 billion in debt, declared bankruptcy fueling the biggest financial crisis since the 1929 great depression. On that date stock markets around the world collapsed and the catastrophic implications of the subprime mortgage lending spread across all global markets. As with the tsunamis practically no one predicted Lehman's default, the 2007/2008 Great Recession, or the global financial crisis that followed [66], instigating the near collapse of the global financial system. Although there have been other serious financial crises in the past (e.g. the 1929 Depression) none was of such massive intensity and worldwide spread. The strong interconnectivity of the global financial markets and the instant transmission of information accelerated the spread of bad news and deepened the depth and spread of the crisis at levels never experienced in the past, causing huge financial losses across the world that exceeded $30 trillion (about the combined GDP of USA, China and UK) and another Black Swan.

In the middle of January 2015 oil prices collapsed to less than $45 a barrel, a huge drop of more than 50% from just a couple of years earlier. Worse still, the great majority of forecasters did not predict such a huge drop raising concerns as to whether there is any value in forecasting when no warnings for such a drop were provided [79]. Oil prices have fluctuated widely in the past. At the end of 1973 they more than doubled in a short period of time, from $4.3 a barrel to beyond $10 causing a serious economic downtrend. Consequently they reached $38 a barrel in January 1981 contributing to what was then called "stagflation", a period of high inflation and no economic growth. Then the price was dropped to $11.5 in July 1986, raised to $40 in September 1990, lowered to less than $15 in May 1992 and increased to more than $100 for most of 2008 and a good part of 2013. Given such huge variations in oil prices during the past, the January 2015 drop was not unusual and should not have caught anyone by surprise. Yet it did, causing considerable economic hardship in oil producing countries like Venezuela that experienced big shortages in basic goods as well as Russia, among others, that saw its currency devalue significantly and its stock market fall. Could the big drop in oil prices have been predicted? The answer is that the exact timing and magnitude of the drop, as with tsunamis, could not have been predicted but the possibility could not have been excluded based on the past behavior of oil prices. This means that the uncertainty and risk could have been evaluated and many things could have been done to have prepared oil producing countries to deal with the consequences of a possible big drop in prices.

In addition to the negative Black Swans there are also positive ones that cannot be predicted and come as a great, pleasant surprise. It is estimated that Jan Koum received more than $8.5 billion from the sale of a company he founded, *WhatsApp*, to Facebook, making him a multibillionaire. Koum, an immigrant from Ukraine, became one of the richest persons in the world with a fortune that at that time was equal to almost 5% of his native country's GDP. Interestingly enough Koum and his co-founder Brian Acton had in 2007 applied for a job at Facebook and they were rejected. Another positive Black Swan took place with the founders of the now giant Google, with a capitalization at the end of 2015 close to $530 billion. At the end of the 1990s,

Google was put up for sale by its founders Larry Page and Sergey Brin, for $1.6 billion and luckily for them there were no buyers at this price. There was only one offer for $750,000 by the portal Excite which Page and Brim refused [9]. The same way Facebook rejected to give a job to Koum and Acton, Yahoo and the other big players of the time denied Google's founders' offer to buy their firm for $1.6 billion. At that time the value of "search" was not obvious thus the future value of Google was seriously underestimated.

Pure luck can also create unexpected surprises. Joan Ginter, for instance, won $5.4 million in a 1993 Lotto Texas draw, and then she won $2 million in Holiday Millionaire in 2006, $3 million in Millions & Millions in 2008, and $10 million in another lottery in 2010, beating all other multi winners. Frane Selak, a Croatian man, was equally lucky in his misfortune. He was one of the few who escaped death in January 1962 when the train he was riding overturned and crashed into a frozen river. In 1963 during his first and only plane ride, he was blown out of a malfunctioning plane door but landed in a haystack avoiding death. In 1966 the bus he was riding skidded off the road and fell into a river and Selak avoided drowning by swimming to the shore. In 1970 and 1973 he missed death in car accidents while in 1995, he was struck by a bus in Zagreb but only sustained minor injuries. Finally, in 2003 he won €800,000 in the lottery. Clearly, luck can produce highly improbable outcomes [40] that cannot by definition be predicted and are not part of what can be considered as forecasting.

The above have illustrated both negative and positive Black Swans that forecasting could not have predicted using conventional approaches. For the negative ones, the critical question is what could have been done (like the establishment of building codes to endure strong earthquakes, or buying insurance policies in case of a fire destroying a home) to assess uncertainty, evaluate the resulting risk and be prepared to cope with their negative consequences. Oil producing countries could, for instance, have signed long term contracts at fixed prices, while Lehman Brothers could have imposed a lower level of leverage and other possible actions, accepting lower profits, for a significant reduction in the risk of bankruptcy (see additional coverage below). Therefore, there are ways to avoid surprises and reduce uncertainty by correctly evaluating the risks involved and taking actions to prepare to manage them. A more general approach is to follow Taleb's [90] suggestions by establishing antifragile strategies, or utilize the methodology of the Extreme Value Theory (EVT) developed in hydrology close to three decades ago [84].

Although there are many situations when accurate forecasting is impossible, there are many millions of others where predictions can provide useful information to improve our decisions and gain from effective action. Weather forecasts, made several times a day, in hundreds of thousands of locations around the world, are a prime example. Such forecasts allow us to prepare for rain, to get appropriately dressed for cold, or to postpone a long trip in case of severe weather conditions and although some of these forecasts may be wrong, on the average their contribution and usefulness is significant and their predictions unbiased. The same is true when companies forecast the amount of inventories they should keep for each of the thousands of products/items they sell to consumers and by doing so improve their effectiveness and profitability. Clearly, some events can be predicted more accurately than others, some cannot be forecasted at all and most are in the in between category. However, what is extremely helpful in all cases is to know what can and cannot be predicted and the extent of uncertainty associated with each type of forecast.

The purpose of this paper is to survey the various disciplines concerned with forecasting in order to assess their degree of predictability and the extent of uncertainty associated with them. In order to do so, the following assumptions are needed:

- Forecasting is based on identifying and estimating, through observation and in some instances theory, patterns and/or relationships and then extrapolating or interpolating them in order to predict. This fact implies four things:
  - There is no other way to forecast the yet unknown future, as no one possesses prophetic powers.
  - In statistical predictions, the estimation of the most appropriate pattern/relationship is done by separating the randomness (noise) from the data. But randomness introduces errors and produces uncertainty in the forecasts.
  - Forecasting errors most of the time behave "normally", but they can also involve "fat" tales and unbounded uncertainty.
  - Patterns and relationships can and do change over time decreasing the accuracy of forecasts and further increasing uncertainty.
- Forecasts can be categorized as:
  - Those intended to identify what is most likely to happen assuming usual conditions, and

Table 1
Forecasting accuracy, uncertainty and risk: From Known/Knowns to Black Swans

	Known	Unknown
Known	I. *Known/Knowns* (majority of real life situations under normal/usual conditions) *Accuracy:* Reasonable (depending on specific factors) *Uncertainty (thin tailed)*: Measurable *Risk*: Can Be Estimated (assuming normality of errors and consistency in the prevailing, normal conditions)	II. *Unknown/Knowns* (knowing but not wanting to believe and act) *Inaccuracy*: Can Be Large (influenced by judgmental biases and irrationality) *Uncertainty*: Large, immeasurable and usually under-estimated significantly *Risk*: Underestimated (due to judgmental biases, irrationality and wishful thinking)
Unknown	III. *Known/Unknowns* (majority of real life situations under unusual/special situations) *Inaccuracy*: Large to Great *Uncertainty (fat tailed)*: Large to Great *Risk*: Hard to Estimate (usually underestimated given the uniqueness of the unusual/special situations)	IV. *Unknown/Unknowns* (Black Swans: Unexpected, surprising events with severe consequences) *Entirely Unpredictable Uncertainty*: Infinite *Risk*: Inconceivable (preparation is possible only by having adopted antifragile strategies)

  - ∗ Those taking place in unusual, or extreme situations like during periods of storms or hurricanes or periods of economic and other crises (booms). In this category the errors have often fat tales and infinite variance.

- The two categories of forecasts require different forecasting methods characterized by thin and fat tailed errors (see below) that influence uncertainty.
- Forecasts are only useful in conjunction with the realistic assessment of their uncertainty, enabling the truthful evaluation of associated risks.

Types of uncertainty. Given the huge uncertainty surrounding all our predictions there is a strong need to incorporate it as an integral part of the overall forecasting effort. Table 1 is way of doing so. It classifies events into the four categories of known/knowns, known/unknowns, unknowns/knowns and unknown/unknowns and summarizes the accuracy and uncertainty in each. There are two conclusions drawn from Table 1. The first is that uncertainty, and therefore risk, varies significantly from one quadrant to another, being thin tailed in quadrant I, unspecified in II, fat in III, and infinite in quadrant IV. The second and probably more important is that uncertainty is not static as it can move from one quadrant to another over time. Such behavior complicates the task of assessing it and becoming prepared to face the eventual, associated risks, topics to be covered later in this article.

2. Weather: The origin of forecasting

Ignoring superstitious attempts, the first efforts to predict the future were made around 650 BC by the Babylonians by observing cloud formations and through them attempting to forecast the weather. Later Aristotle, at around 340 BC, discussed weather in his book *Meteorological* [3] while his student Theophrastus of Eresus [95] wrote the first weather forecasting book *On Winds and On Weather Signs*. At about the same time Chinese had divided the year into 24 periods, each associated with a different type of weather. Although, most attempts to predict the weather relied partly on observation, they were mostly based on folklore. Nothing much changed until the Middle Ages when the thermometer (1593) and the barometer (1643) were discovered allowing for the more precise measurements of atmospheric conditions and heralding the beginning of a new era of more precise information and more reliable forecasts. Weather forecasting, as all other types is based on observing and identifying weather patterns to figure out the seasons of the year, the variations in temperatures during the day and night and the raining periods among others. Discovering and estimating such patterns forms the foundation of all weather predictions and is responsible for its improved accuracy and higher reliability during the last several decades.

The introduction in the 1920s of the "numerical weather prediction" approach [56] provided a turning point by utilizing mathematical models of current atmospheric and ocean conditions to predict future ones. This type of prediction, however, required heavy computations[1] and could not be applied until fast computers became available in the 1960s. Since then weather forecasts have improved considerably through a combination of the following three factors:

[1] It took Richardson six weeks to produce by hand a six-hour forecast for the state of the atmosphere over two points in central Europe.

- Improved observational and measurement techniques permitting a more accurate representation of the current state of the atmosphere and oceans.
- Advances in available weather forecasting models allowing the incorporation of greater complexity and more sophistication.
- Faster computers capable of running the more complex/sophisticated models and being capable of doing so at shorter time intervals.

The above have improved the accuracy of predictions while forecasts are issued for progressively narrower geographical areas and for shorter intervals of time expanding their value.

There are still limits to accurate weather predictions. In an article published in 1963, Edward Lorenz [52], the father of the chaos theory, suggested that because of the prevailing chaotic atmospheric conditions, there are practical limits to accurate weather predictions, making them impossible beyond two to three weeks. The current, confirmed limit with a more than 60% success rate, over naïve (persistent) predictions is around ten days while research efforts in order to improve such limits are under way. Accepting Lorenz's limit of a few weeks raises serious questions about the appropriateness of long term climate predictions and the claims of global warming while it is well known from Antarctic Core Data that there are huge cyclical variations in the earth's temperatures.

More severe limits exist for the accurate predictions of extreme weather events (severe storms, hurricanes, tornados, intense snowfalls, heavy rains and floods, and extended periods of droughts) caused by lack of understanding of the complex physics characterizing these events that are quite different to those of the usual weather conditions. At present considerable research efforts are being undertaken to better understand these extreme conditions and to be able to improve their predictability, or at least develop warning signals of impending severe weather conditions as early as possible.

Weather forecasting started as folklore and slowly has been turning into a solid scientific discipline. It evaluates objectively the accuracy of its predictions and uses feedback to further improve them. Moreover, meteorologists are well aware of the uncertainty in their predictions, providing probabilistic forecasts, and are constantly searching for enhancements using objective feedback. The uncertainty of weather forecasts is well calibrated [77] increasing their everyday value as people are learning to trust them. This does not mean that they are always accurate, however, their overall success rate is improving and their assessment of uncertainty is well calibrated. Finally, weather forecasters have learned that predicting extreme weather events requires different models and skills than those of normal ones [12]. At the same time, they consider such events as an integral part of their job, even though it requires special effort, different models and extra skills to predict them.

Next the various types of forecasting are surveyed under two broad headings. First, those that involve *natural phenomena* where humans cannot influence their future course, except to a limited extent, and second the great majority of events in *the social world* that humans can and do influence with their actions and reactions, changing their future course, making forecasting more difficult but also more challenging.

3. Forecasting natural events

Eclipses of the sun or the moon, the trajectory of space ships travelling beyond our solar system, the path of comet Harley and the prediction that it will approach the earth again on July 21, 2061, the ability of a skyscraper to withstand earthquakes of certain magnitude as well as a great number of similar events following some natural laws or engineering principles allow for practically perfect predictions and therefore encompass uncertainty approaching zero. But most others like earthquakes, tsunamis, volcanic eruptions, floods and so on cannot be predicted, sometimes at all, involving a great amount of uncertainty. In addition to weather forecasting that was already covered, there is below a brief survey of our ability to forecast various natural events and the uncertainty involved in doing so.

3.1. Earthquakes

The mechanism generating earthquakes is activated when two plates of the earth suddenly slip past one another on a geological fault. It is the sudden release of energy that causes a seismic wave that shakes the area around the epicenter of the earthquake. Earthquakes of high magnitude are rare while those of lower ones occur frequently on a regular basis. The distribution of earthquakes follows an almost perfect power law (see Table 2).

The regularity of Table 2 refers to the whole world and covers the average of an entire year. But regularity does not indicate predictability. Unfortunately, the exact timing, specific location as well as the precise magnitude of an earthquake is totally unpredictable. Sci-

entists have tested many different theories and possible indicators that could predict earthquakes. There are claims of strange animal behavior before earthquakes [97] or fluctuations in underground activity [98] that could be used as early indicators of forthcoming ones. But so far no proof of success has been established. Scientists can predict the general location where major earthquakes are likely to occur based on the existence of fault regions. They can also make probabilistic estimates of when a certain earthquake may strike, by considering the past history of earthquakes in that region. But such predictions are probabilistic, referring to the long term and based on the average, and cannot predict specific, future ones. At the same time scientists have been more successful in predicting aftershocks and in guessing how an earthquake at one point along a fault may affect new ones in interconnected faults.

Table 2
Frequency of occurrence of earthquakes globally*

Magnitude	Average (annual)
8 and higher	1
7–7.9	15
6–6.9	134
5–5.9	1,319
4–4.9	13,000*
3–3.9	130,000*
2–2.9	1,300,000*

*Estimated. *Source*: USGS http://earthquake.usgs.gov/earthquakes/eqarchives/year/eqstats.php.

The history of earthquake predictions is one of complete failure. Richter [80] the developer of the magnitude scale for earthquakes, commented: *"Journalists and the general public rush to any suggestion of earthquake prediction like hogs toward a full trough... [Prediction] provides a happy hunting ground for amateurs, cranks, and outright publicity-seeking fakers"*. This comment still holds true according to Geller et al. [29] who adds:

> *"Earthquake prediction is usually defined as the specification of the time, location, and magnitude of a future earthquake within stated limits... Previous Perspectives in Science may have given a favorable impression of prediction research, and the news media and some optimistic scientists encourage the belief that earthquakes can be predicted. Recent research suggests to us that this belief is incorrect"* ([29], p. 1616).

The unpredictability and uncertainty of earthquakes can be seen in Table 3 that lists the seventeen largest earthquakes of the 20th and 21st century, with over 100,000 deaths, by magnitude.

In terms of Table 1 earthquakes fall into the "Known/Unknowns" or "Unknown/Unknowns" categories excluding attempts to accurately predict them. Lack of predictability increases uncertainty and raises

Table 3
The seventeen largest earthquakes in the World since 1900

	Location	Date UTC	Magnitude	Latitude	Longitude
1	Chile	1960 05 22	9.5	−38.29	−73.05
2	1964 Great Alaska Earthquake	1964 03 28	9.2	61.02	−147.65
3	Off the West Coast of Northern Sumatra	2004 12 26	9.1	3.3	95.78
4	Near the East Coast of Honshu, Japan	2011 03 11	9	38.322	142.369
5	Kamchatka	1952 11 04	9	52.76	160.06
6	Offshore Maule, Chile	2010 02 27	8.8	−35.846	−72.719
7	Off the Coast of Ecuador	1906 01 31	8.8	1	−81.5
8	Rat Islands, Alaska	1965 02 04	8.7	51.21	178.5
9	Northern Sumatra, Indonesia	2005 03 28	8.6	2.08	97.01
10	Assam–Tibet	1950 08 15	8.6	28.5	96.5
11	Off the west coast of northern Sumatra	2012 04 11	8.6	2.311	93.063
12	Andreanof Islands, Alaska	1957 03 09	8.6	51.56	−175.39
13	Southern Sumatra, Indonesia	2007 09 12	8.5	−4.438	101.367
14	Banda Sea, Indonesia	1938 02 01	8.5	−5.05	131.62
15	Kamchatka	1923 02 03	8.5	54	161
16	Chile–Argentina border	1922 11 11	8.5	−28.55	−70.5
17	Kuril Islands	1963 10 13	8.5	44.9	149.6

Source: http://earthquake.usgs.gov/earthquakes/world/10_largest_world.php. Updated 2012 April 11.

the question of what can be done to at least moderate their negative consequences, as earthquakes can kill a great number of people each year and cause billions of dollars in infrastructure and property damages. In high risk countries like Japan and regions like Southern California the emphasis shifts from forecasting to preparation. This is done by improving the ability of buildings to withstand earthquakes (balancing the costs and the maximum expected earthquake magnitude), reducing the chance of fires after earthquakes, minimizing the destruction in infrastructure as well as the number of deaths. Another approach is based on conducting drills to prepare people on how to respond in case of an earthquake.

But there is still uncertainty that cannot be eliminated leading to Black Swans and the fourth quadrant of Table 1. The 2011 magnitude 9.0 earthquake at Tohoku, Japan (triggering the catastrophic tsunami that killed an estimated 29,000 people) was four times stronger than the previous largest one of 8.6 recorded in 1707. Thus, the claim by the US Geological Society [96] that the probability of a major earthquake occurring in the San Francisco Bay area over the next 30 years is 67% does not take into account that the next earthquake can be stronger or earlier than previous ones in the Bay area. In addition, it does not provide a range of uncertainty around the "30 years" estimate or the "67% probability". Unfortunately, the possibility of Black Swans always exists, and by definition they cannot be predicted making the USGS estimates based on past information alone misleading, providing a false sense of security. As the case of Tohoku in Japan is a fact that cannot be ignored making it necessary to broaden the historical information to include pre-historical geological evidence, or to extrapolate Table 3 to encompass much larger earthquakes that are possible using basic physical laws and EVT (the alternative of searching for broader evidence applies to the other forecasting problems, too).

3.2. Tsunamis

Tsunamis are enormous waves usually created by underwater disturbances when large areas of the sea floor are displaced, mainly because of large earthquakes and in some cases by strong volcanic eruptions, landslides or even asteroids falling in the ocean. As the prediction of earthquakes is not possible so is that of tsunamis. In addition, tsunamis, at least big, catastrophic ones, are rare events and historical data of their appearance are few making the estimation of uncertainty associated with them practically impossible. With Tsunamis the emphasis has shifted from forecasting them to monitoring their trajectory, once a big earthquake has occurred in or near the sea, in order to warn coastal areas of their impending arrival. Monitoring is based on a network of sensors in the sea and through satellites to detect them, using a communications infrastructure to issue timely alarms for the evacuation of coastal areas ([30], p. 245).

Several projects are under way to detect tsunamis as soon as they occur, using special devices anchored at the bottom of oceans in tsunamigenic prone zones. The hope is that the right usage of these signals will provide accurate information on the direction and the magnitude of tsunamis so that early warning signals can be issued in the affected areas. However, these systems are still under "research in progress" with no certainty that they will correctly detect forthcoming tsunamis. At the same time after large earthquakes there are many false tsunami alarms recommending unnecessary evacuations. As a consequence people do not believe the warnings when real ones occur.

The uncertainty caused by tsunamis, and probably other natural disasters, can be appreciated in Table 4. Before 2004 the deadliest one was in Lisbon in 1755 causing 60,000 casualties. The Sumatra tsunami, on December 26, 2004, left almost four times as many dead. In addition, there were only seven years between the 2004 and the 2011 tsunamis while in the past the interval between large ones was many times more. Finally, the 2011 tsunami caused a serious nuclear accident at three reactors in the Fukushima Daiichi Nuclear Power Plant complex that necessitated the evacuation of hundreds of thousands of people, further raising the uncertainty of the extent of possible damages to be caused by future ones. As with earthquakes, tsunamis fall into the third and fourth quadrant of Table 1, excluding attempts of accurate prediction and leading to huge uncertainty and great risks.

3.3. Volcanos

During the last decade significant progress has been made to better understand the mechanisms that trigger lava-forming and explosive volcanic eruptions. Yet, the predictability of volcanic eruptions remains low even after signal devices aimed at detecting seismic activity in the area have been installed and new satellite observations have been used. As with tsunamis the emphasis has shifted from attempting to predict longer term volcanic eruptions to monitoring their short term (few

Table 4
The ten worst tsunamis

Place	Occurrence	Number of dead	Max. wave height (m)	Magnitude of earthquake
Sumatra, Indonesia	26 December 2004	230,000	50	9.1
North Pacific Coast, Japan	11 March 2011	18,000	10	9.0
Lisbon, Portugal	1 November 1755	60,000	30	8.5
Krakatau, Indonesia	27 August 1883	40,000	37	Volcano eruption
Enshunada Sea, Japan	20 September 1498	31,000	8.3	8.3
Nankaido, Japan	28 October 1707	30,000	10	8.4
Sanriku, Japan	15 June 1896	22,000	38.2	7.6
Northern Chile	13 August 1868	25,000	21	8.5
Ryukyu Islands, Japan	24 April 1771	12,000	85.4	7.4
Ise Bay, Japan	18 January 1586	8,000	6	8.2

days to a maximum of a couple of weeks) behavior. Another approach involves the use of time series of past volcanic eruptions in order to predict future ones currently under way through the usage of analogies. In some models being currently developed the input variable is the local seismicity rate, e.g. its daily average, while the output is the actual eruption in say five or six days in the future. Other models strive to predict longer term eruptions by associating earthquake information with various types of potential eruptions. Although forecasting remains the main aim of volcanology progress is still elusive.

The Volcanic Explosivity Index (VEI), similar to the Richter scale for earthquakes, measures the relative explosiveness of eruptions, how much volcanic material is ejected, the height of the material thrown and how long the eruptions last. As with earthquakes, volcanic eruptions with an index greater than 4 are exceptionally rare to highly infrequent, making any long term prediction highly uncertain. Yet when they occur they can cause havoc in the surrounding areas as well as in air traffic that cannot operate in skies covered with volcanic ash under the fear it can damage airplane engines. Although volcanic explosions, like earthquakes and tsunamis fall into the third and fourth quadrat, the uncertainty and risk involved can be assessed easier as volcanos move slowly and there are plenty of warning signs reducing the likelihood of Black Swans.

3.4. *Floods*

Floods occur when water inundates land that is normally dry. Floods may happen from heavy rainfall, overflow of water from rivers or lakes, intense sea tides and/or when water overtops or breaks levees or dams. Floods can cause damages, sometimes serious, to homes, farmland and animals, businesses as well as infrastructures. Some floods develop slowly, while others can grow in minutes because of heavy rainfall or other causes. Floods can be local or can affect large areas usually around rivers or in flatlands. The predictability of floods is better than that of the other natural calamities surveyed so far. Several methods have been developed to predict floods on a real time basis as a function of rainfall or snow melting with the purpose of providing early warnings signals. The major objective of early warnings is taking actions to minimize the extent of floods by diverting river water to uninhabited areas, or to provide help to flood victims as early as possible in order to reduce the loss of life and the hardship to those being affected. As with the other natural disasters described so far, floods fall into the third and fourth quadrat of Table 1 with devastating consequences and possible Black Swans creating turmoil and huge damages to the affected areas. Hydrology is, however, an area that has made progress in estimating uncertainty using the EVT.

3.5. *Other natural disasters (hurricanes, tornados, droughts, forest fires, avalanches)*

In addition to those already mentioned, there are additional natural disasters that can cause serious damage in terms of both death and property/infrastructure destruction. The ability to predict these disasters is limited or non-existent, shifting prevention efforts to the most effective monitoring possible to warn of their impending arrival as in the case of hurricanes and tornados. Alternatively, efforts are being concentrated to prevent them by continuous watching as in the case of forest fires, in order to extinguish them as early as possible, or to avoid their negative consequences as with avalanches, which in a number of cases can be initiated with explosives before they become danger-

ous. Droughts can in some instances be anticipated and steps can be taken to minimize their negative consequences, but some extensive, serious ones are due to lack of rain for long periods of time and are hard to predict while their harmful effects are tougher to deal with.

Natural disasters fall into the first quadrant the great majority of the time as there are no earthquakes, tsunamis, floods, hurricanes and so forth. The problem is when such events occur pushing them into the third and fourth quadrant of Table 1 with devastating consequences as Black Swans are possible.

3.6. Climate change

Over the last couple of decades there has been a heated debate about the risk of dangerous global warming arising from carbon dioxide emissions from human activity. The U.N.'s Intergovernmental Panel on Climate Change (IPCC) projects that unless carbon dioxide emissions are severely curtailed, harmful warming will occur. Skeptics, however, point out that the physical-science assumptions behind the projections of dangerous warming are doubtful [99] and that the IPCC's long-term trend projections of temperatures and environmental effects from the short-term behavior of temperatures are inconsistent with good forecasting practice and are less accurate than the no-change forecast [37]. Evidence collected during the past 30 years indicates that the Earth has a moderate, persistent 1,500-year climate cycle that creates warmings and coolings. Such cycles exist in the many graphs based on Antarctica and Greenland ice core data (see [33]).

In a study sponsored by the National Research Council in the USA [78] it was concluded:

> *"Large, abrupt climate changes have affected hemispheric to global regions repeatedly, as shown by numerous paleoclimate records. Changes of up to 16°C and a factor of 2 in precipitation have occurred in some places in periods as short as decades to years"* (p. 10).

Given such evidence and the existence of long term temperature cycles, can global warming be justified from data showing temperature increases in a single century? Unfortunately, there are strong arguments from those claiming that there is no objective evidence proving global warming and those claiming that the evidence is clear that our Earth is warming and that we, humans, are largely responsible for that [20]. McGregor is proposing an enhanced role of climatology to collect and make available relevant information to aid the environmental risk management and avoid the polarization between the environmentalists and those opposing any action [71]. The almost religious attitudes of the two sides increases uncertainty and point to the idea that "the burden of evidence is on those who disturb natural systems" [90,92].

4. Forecasting social events

The previous section on forecasting natural events indicated the great difficulties and in the great majority of cases the impossibility of accurate predictions of the specific location, the precise time of occurrence of the event and its magnitude. This section surveys forecasting in various areas of social science and discusses the issuing uncertainty that varies widely from one area to another.

4.1. Economy and business

The areas of economy and business have witnessed, more than any other discipline, the greatest effort in developing the widest variety of forecasting methods and models, in producing the biggest number of predictions and in employing the largest number of forecasters. In the process there have been huge failures but also some concrete successes. After a short history of forecasting, the major purpose of this paper is to review both successes and failures, discuss the reasons involved and recommend specific actions to improve the field.

A short history of economic and business forecasting. Quantitative economic and business forecasting has developed along the two distinct categories of time series and econometric. A survey paper [59] chronicles the history of time series predictions as follows:

> *"Time-series considerations originated in 1807, when the French mathematician Fourier claimed that any series can be approximated as the sum of sine and cosine terms. This idea was used by Schuster (1906) who applied Fourier's expansion to estimate the length of hidden periodicities and who widely utilized periodogram analysis in his work. The modern era in time series started in 1927 with Yule and achieved its major advances in 1938 when Wold developed a comprehensive theory of Autoregressive/Moving Averages (ARMA) schemes, around 1940 when Wiener and Kolmogoroff solved the estimation problem for continu-*

ous and discrete filters correspondingly, and in the early '60s when Kalman and Kalman and Bucy extended Wiener and Kolmogoroff's estimation procedures to non-stationary series involving systems in the time domain. On the Operation-Research side, the late '50s and early '60s saw the development of exponential smoothing models which, simple as they were, became utilized by business firms and the military. In the area of decomposition, the utilization of digital computers opened a new era by allowing the cumbersome computations, done beforehand on a desk calculator, to be easily performed by the computer" (pp. 29–30).

In an updated paper, Gooijer and Hyndman [34] have reviewed the past 25 years of developments into time series forecasting and commented on directions for future research. They pointed out that although enormous progress has been made, there are still a large number of topics needing further development. Along similar lines Armstrong and Fildes [5] describe the advancements in the field and stress the need for *evidence based forecasting* to guide practitioners and researchers and the need to avoid resistance from both groups when the findings do not agree with their prior beliefs and/or personal interests. Furthermore, Armstrong, Green and Graefe [6] propose *conservatism* as the golden rule of forecasting and provide evidence that adopting this simple guideline (compared to common practice) reduces errors by 28%. Finally, Green and Armstrong [38] in a special issue of the *Journal of Business Research* on "Simple Versus Complex Forecasting" have shown that complexity does not improve predictive accuracy.

On the econometric side the method of regression was discovered almost concurrently with that of time series when Legendre, in 1805, and Gauss in 1809 applied the method of least squares to determine astronomical relationships. The term "regression" was conceived by Galton in 1886 when he showed that the heights of descendants of tall parents tended to regress down towards the average height (a regression towards the mean). Consequently, his work was extended by Yule in 1893 and Pearson ten years later to include any relationship and was further improved by Fisher in the early to mid-1920s. In the early 1950s regression became practical through the utilization of desk calculators while afterwards digital computers allowed its widespread utilization to identify and estimate relationships between a dependent and one or more independent ones in a single equation. Econometrics extended regression by simultaneously including more than one equation into the model whose parameters were estimated simultaneously using the method of least squares as well as other estimation procedures. The utilization of regression and econometric models was extended by its adoption, in addition to governmental organizations, by mostly large business firms.

Regression and econometric models cannot provide forecasts directly like their equivalent time series ones as they require the prediction of future values of the independent variables being used as well as judgmental inputs about fiscal and monetary policies, global competition and a host of other factors. Yet they were used considerably in the past, although such efforts have been reduced substantially in the economic sphere and have been completely abandoned by business firms. The reason is simple. The predictive accuracy of econometric forecasting has not matched expectations. A paper by Armstrong [4] "Forecasting with econometric methods: Folklore versus fact" describes the empirical evidence and comes up with the following two conclusions that still hold true today:

- Simpler econometric models were found to be more accurate than complex, sophisticated ones; and
- Simple, mechanical time series models were found to be more accurate than econometric ones.

In a recent survey, Green and Armstrong [38] found that simple causal models that represent prior knowledge provide forecasts that are more accurate than those from complex statistically estimated econometric models. In another paper Allen and Morzuch [2] describe the progress, problems, and conflicting evidence in econometric forecasting and the lack of concrete improvements in terms of forecasting accuracy. Statisticians claim there are cases where multiple regression models can improve predictive accuracy but do not explain how the independent variables are extrapolated and what their contribution is in improving accuracy. Econometricians insist on the usefulness of their models, arguing that they are able to explain economic relationships and provide useful insights for policy makers, making their poor predictive accuracy irrelevant. In a special issue of the *Journal of Econometrics* [41] the emphasis is "in the progress of building more sophisticated models", with not a word on empirical evaluations and the predictive accuracy of such models.

A more severe critique of econometric models comes from Lucas [54]:

"Given that the structure of an econometric model consists of optimal decision rules of economic

agents, and that optimal decision rules vary systematically with changes in the structure of series relevant to the decision maker, it follows that any change in policy will systematically alter the structure of econometric models".

Lucas critique, or law, is along the same lines with Campbell's and Goodhart's law which state that the more a measure is used the greater the chance that it will distort and corrupt the social processes it intends to monitor.

On the other hand, the predictive accuracy of time series methods has been tested extensively over the last four decades in a large number of empirical studies. The experimental design of such testing has replicated reality and left little doubt about their correctness. The conclusions of these studies are the following:

- In actual post-sample comparisons, simple, mechanical models such as exponential smoothing or even the benchmark method of random walk, or its seasonal equivalent, outperformed, on average, statistically sophisticated ones such as ARIMA models, Kalman filters, Bayesian methods, neural networks and expert systems, that require considerable expertise to model and ample computer time to run.
- Statistically sophisticated methods were considerably more accurate in the model "fitting" phase (that is explaining what has happened during the past) without such superiority to translate to better predictions. Makridakis and Winkler [69] found that the R^2 between model fit and forecast performance was 4% for the first three forecasting horizons, dropping to 1% for period 5 and 0 and for period 12 and beyond.
- Contrary to statistical theory it was found that a larger sample size did not improve the size of the forecasting errors or reduce the confidence intervals of forecasts [55,61,63].

As the bulk of time series forecasting applications have been in the area of business the above conclusions have been adopted and have significantly influenced the practice of forecasting among business firms. At the same time theoretical statisticians have largely ignored the empirical evidence emphasizing the development of more sophisticated models even though such models do not outperform simple ones [23].

Why simple forecasting models outperform sophisticated ones. We have mentioned that forecasting is founded on the identification and extrapolation of established patterns to predict their continuation. But apart from Newtonian mechanics and similar exceptions, the future is never exactly like the past which means that the accuracy of extrapolative predictions cannot be assured to be accurate. The crucial question is the extent of accuracy, or inaccuracy of such predictions. The problem is that most of the time series in the economic/business world are influenced by random events and often behave not far from random walks, favoring simple methods that are capable of smoothing such randomness. The forecasting ability of simple time series methods is based on the following two factors widely used in weather predictions:

- *Persistence in immediate term predictions*: In weather, forecasting that tomorrow's conditions will be the same as those of today are pretty accurate as there is some impetus in weather patterns which in some places like Southern California or Athens in the summer can last for many weeks and even months. In others like London or Paris the persistence of such impetus is much shorter while in the great majority of places lasts for a few days and is greatly exploited by forecasters to predict weather conditions for up to three days in advance. The persistence exploited in weather predictions is called momentum in economic and business patterns and is the major factor influencing their accuracy. However, such momentum deteriorates with time even more rapidly than that of weather as people's actions and external events can and do influence economic/business patterns to a greater extent than those of weather. Simple forecasting methods by "averaging" past errors capture the immediate term momentum more accurately than sophisticated ones that attempt to discover more elaborate but often, illusive patterns.
- *Meteorology in short term, seasonal predictions*: The average conditions over the last usually 20 years, is another way employed in weather predictions. Such an approach has been found to be superior to other alternatives and it is also used by business and economic forecasters to make daily, weekly, monthly or quarterly seasonal predictions utilizing the consistency of repetitive seasonal patterns. Moreover, empirical studies have found that simple averaging of seasonal patterns, after the extreme high and low of each season has been eliminated, provides more accurate predictions than more complicated alternatives. Simple methods can, therefore, capture the seasonal pat-

tern in the data accurately and extrapolate it to improve predictive accuracy.

Forecasting and uncertainty beyond the immediate and short term. Beyond the immediate and short term the accuracy of predictions drops while uncertainty increases. A major reason is cyclical fluctuations that cannot be predicted as their duration and depth varies widely from one cycle to another [53] The difficulty in forecasting cycles can be seen in Table 5 that lists all major ones since 1900 classified as contractions (recessions) and expansions (booms). For instance the length of recessions, from peak to trough, varies considerably from 6 to 43 months while that of expansions from 10 to 120 months. Moreover, their depth varies widely as can be confirmed by the drop in GDP that was −2.2% in the 1980 recession to more than −34% in the 1929/1933 one. Also the rates of growth differ widely, ranging from 4.4% in the 1980/81 expansion to 41.0% in the 1991/2001 one. It is important to underline that such numbers refer to the aggregate of GDP and fluctuate much more at less aggregate levels as those referring to industries, individual firms not to mention individual products and even single items.

Similar fluctuations exist during periods of other crises, or booms as during the 1973 oil embargo that created an abrupt financial crisis, or the late 1990s dot-com euphoria (or "internet bubble") that saw the NASDAQ index rising to the historic height of 5,049 in

Table 5

National Bureau of Economic Reseaech (NBER): Length of business cycles

Business cycle		Duration in months			
Quarterly dates are in parentheses		Peak to trough	Previous trough to this peak	Trough from previous through	Peak from previous peak
	December 1900 (IV)				
September 1902 (IV)	August 1904 (III)	23	21	44	39
May 1907 (II)	June 1908 (II)	13	33	46	56
January 1910 (I)	January 1912 (IV)	24	19	43	32
January 1913 (I)	December 1914 (IV)	23	12	35	36
August 1918 (III)	March 1919 (I)	7	44	51	67
January 1920 (I)	July 1921 (III)	18	10	28	17
May 1923 (II)	July 1924 (III)	14	22	36	40
October 1926 (III)	November 1927 (IV)	13	27	40	41
August 1929 (III)	March 1933 (I)	43	21	64	34
May 1937 (II)	June 1938 (II)	13	50	63	93
February 1945 (I)	October 1945 (IV)	8	80	88	93
November 1948 (IV)	October 1949 (IV)	11	37	48	45
July 1953 (II)	May 1954 (II)	10	45	55	56
August 1957 (III)	April 1958 (II)	8	39	47	49
April 1960 (II)	February 1961 (I)	10	24	34	32
December 1969 (IV)	November 1970 (IV)	11	106	117	116
November 1973 (IV)	March 1975 (I)	16	36	52	47
January 1980 (I)	July 1980 (III)	6	58	64	74
July 1981 (III)	November 1982 (IV)	16	12	28	18
July 1990 (III)	March 1991(I)	8	92	100	108
March 2001 (I)	November 2001 (IV)	8	120	128	128
December 2007 (IV)	June 2009 (II)	18	73	91	81
June 2009 (II)	January 2015		**55**		
	Min	**6**	**10**	**28**	**17**
	Max	**43**	**120**	**128**	**128**
Average, all cycles (since 1854):					
1854–2015 (33 cycles)		17.5	38.7	56.2	56.4
1854–1919 (16 cycles)		21.6	26.6	48.2	48.9
1919–1945 (6 cycles)		18.2	35	53.2	53.0
1945–2009 (11 cycles)		11.1	58.4	69.5	68.5

March 10, 2000, and then falling to more than half on January 2, 2001 producing, first a boom and then a recession in the USA and global markets (actually, more than fifteen years later, on June 16, 2015 the NASDAQ index reached the same value to that of March 10, 2000). The conclusion is that crises or booms do not follow constant patterns, or exact relationships, generating unique conditions that cannot be predicted beforehand, thus, requiring a new approach and novel thinking to face them. In terms of uncertainty forecasts move from quadrat I to III and the possibility of even IV as Black Swans can also develop as during the 1929/33 Depression or the 2008/09 Great Recession.

Economic and business forecasters must follow the example of their colleagues in weather who utilize separate approaches and different models to predict normal weather conditions and unusual ones, as the atmospheric conditions are quite different in the two situations. It must be clear that established patterns/relationships, determined during normal conditions, can only be extrapolated (or interpolated) during such conditions and cannot be applied in unsettled periods as those prevailing during recessions/booms or other crises/good times, a fact that can be confirmed by the disastrous track record of economic and business predictions [64]. The current practice in the areas of economy and business is to label as "outliers" the situations falling outside normal conditions and ignore them. Weather forecasters, on the other hand, utilized different models and change their emphasis from more accurate predictions to better monitoring by following as closely as possible the progress of the new, developing weather patterns. In the case of hurricanes, for instance, although their exact location and timing cannot be predicted in advance, once one has been identified it is closely monitored to figure out its course and strength as accurately as possible, thus providing warnings, as early as possible, to all likely areas to be affected.

In the area of inventory management there is a clear distinction between normal forecasting of demand and that of what is called intermittent one [88] as in the case of a machine breakdown when parts to repay it are needed. In such cases the demand is zero for long periods of time and then peaks with the failure. Intermittent demand forecasting has explored and proposed the most appropriate modes. However, as Gooijer and Hyndman [34] conclude, it is surprising that so little work has been done in this area. In addition, forecasting when special events/actions are involved uses a distinct approach as they modify the established patterns or relationships and require to figure out their forecasting influence.

Stock (and other) markets. Accurate predictions of stocks and other markets can lead to vast profits if they turn out to be correct and for this reason there is a great number of "how to do" books promising success and riches. Empirical evidence, however, has proved that such advice is worthless. Markets are efficient [21,70] making their accurate predictions impossible, apart from having inside knowledge, an illegal activity, or being able to gain a technological advantage over others as with the use of high frequency trading [51].

Unlike weather predictions or those in the economic and business world, stock and other markets are different, characterized by the following three factors:

Long term consistency. In advanced economies and at the aggregate level, stock markets increase at a constant real (that is above inflation) rate over the long term, as it is confirmed with data going as far back as 1800. In the USA for instance the stock market has been increasing at the real rate of more than 6.5% a year between 1800s and 2015, while those of other countries are growing at rates ranging from 1.8% for Belgium to 7.4% for Australia ([17], pp. 61–63). Predictions made by books [31] like "Dow 36,000" claiming big improvements have failed miserably the same way as prophesies of doom Batra [8]. On the other side, there are no assurances that long term trends will continue into the future as has been the case of Japan whose Nikkei was increasing consistently in the 1960s, 1970s and 1980s reaching close to the 40,000 mark before the beginning of 1990. But then it reversed its trend dropping to around 8,000 in April 2003 and being below 20,000 for most of 2015.

The same long term consistency applies in commodity markets where real prices decrease exponentially over the long run [82,83] but which exhibit huge cyclical fluctuations, similar to those of stocks, around such a long term trend making their medium term prediction problematic ([65], Chapter 7).

Huge medium term cyclicality. The variability of cyclical fluctuations in the stock and commodity markets is usually significantly greater than those in the economy mentioned above. For instance, during the 2007/8 Great Recession while USA's GDP fell by less than 8% that of the stock market (S&P500) dropped by a little more than 50%. The cyclical fluctuations are wider on industries and even larger on individual stocks, particularly those influenced by cyclical factors. For instance, the NASDAQ index, mentioned above, just recovered its height of 5,049 of March 10,

2000 close to fifteen years later. It is the medium term, cyclical volatility that renders stock market forecasting so difficult and challenging, making the entry and exit points critical for profits or losses when investing.

Short and immediate term momentum. The short and immediate term predictions for the stock and commodity markets as a whole as well as that of individual shares is their latest available price (what weather forecasters call persistence and business ones naïve predictions). That is they behave in a random walk fashion. Available evidence indicates that it is practically impossible to beat the random walk model as the market is efficient and discounts any good or bad news immediately affecting markets as a whole or those of individual stocks/commodities.

4.2. Uncertainty in the stock and other markets

The uncertainty in the stock and other markets is like that of weather. During normal economic and business conditions, markets operate in the first quadrat of known/knowns, with uncertainty being measurable and fluctuating within expected limits in which case risk can be estimated and controlled. But like weather during periods of storms and hurricanes, markets become unpredictable and highly volatile during periods of economic and other crises (or booms) exhibiting fat tailed errors, moving them to the third and even the fourth quadrat as Black Swans are also possible. Unlike weather, however, uncertainty in stock and other markets intensifies considerably during such periods driven by the "animal spirit" of fear and greed as people attempt to minimize their losses or maximize their gains. Thus, psychological factors are in addition significantly influencing market movements, as people seem to overreact to both good and bad news, increasing uncertainty and risk. During such periods price movements do not follow normal distributions but power laws, with extremely small (or large) values many times below (or above) the historical standard deviation. As with weather the forecasting and the assessment of uncertainty of stock and other markets cannot be based on the usual forecasting models and alternatives must be considered as those recommended by Taleb [89] in his four quadrant model that takes into account complex payoffs as well as two distinct types of randomness.

The Chicago Board Options Exchange (CBOE) has introduced an exchange traded Volatility Index (VIX), also known as the "fear" index, measuring the expected uncertainty (volatility) in the S&P index over the next 30 days. Although, there are several problems with the VIX index, its biggest advantages are that it can be used as a means of diversification, as it correlates negatively with stock market returns, and as a way of measuring uncertainty at all times, including recessions and other economic crises (or booms).

4.3. Judgmental forecasting

Judgmental forecasts are the only alternative when no quantitative data is available. A more common situation, however, is to modify the statistical forecasts by incorporating additional judgmental inputs, or inside knowledge not contained in the available data. Judgmental forecasts are much more expensive to prepare than statistical ones and become practically impossible when large numbers of predictions are required. But their high cost is not their biggest drawback. Judgmental predictions suffer from systematic biases that decrease, sometimes substantially, their accuracy. The big challenge is, therefore, to use judgment because of its uniqueness and the extra information/knowledge it possesses but, at the same time, avoid or minimize its biases. Considerable research has shown that mechanical, statistical forecasts are more accurate than judgmental ones [73]. A prime example is the stock market where actively managed funds cannot beat index funds (called ETFs, Exchange Traded Funds) that follow market indexes selecting stocks randomly. In a recent study that included 2,862 actively managed funds conducted by the S&P Dow Jones team, it was concluded that *"most people shouldn't even try to beat the market: Just pick a low-cost index fund, assemble a balanced and appropriate portfolio for your specific needs, and give up on active fund management"* [94]. A similar conclusion has been found by other studies analyzing the success of investment newsletters [36], the recommendations of investment gurus [46], or the judgmental forecasts of companies [49] that research found to be overly optimistic, raising the average error from 32% to 65% [22]. Such findings are raising concerns about the extent of biases of human forecasters as their accuracy is usually lower than that of simple statistical ones lacking inside information and special knowledge, the most important problem being the judgmental bias of overoptimism.

Table 6, taken from Makridakis [60] lists the most relevant of judgmental biases affecting forecasting and indicates ways of avoiding or minimizing their negative consequences. The problem is that people are not

Table 6
Common judgemental biases and ways of avoiding or reducing their negative impact

Type of bias	Description of bias	Avoiding/reducing negative impact
Search for supportive evidence	Willingness to gather evidence that lead toward certain desired conclusions and to disregard threatening evidence	• Encourage disconfirming evidence • Introduce devil's advocate(s)
Inconsistency	Inability to apply the same decision criteria in similar situations	• Formalize decision making • Create decision making rules
Conservatism	Failure to change (or changing slowly) one's own mind in light of new information/evidence	• Monitor changes and build procedures to take actions when such changes are identified
Recency	The most recent events dominate to a greater extend those in the less recent past, which are downgraded or ignored	• Accept the existence of cycles that means that not all ups or downs are permanent • Consider the fundamental factors affecting events of interest
Availability	Reliance upon specific events easily recalled from memory, to the exclusion of other pertinent information	• Consider all available information • Utilize both recent and past information that covers all sides of the argument
Anchoring	Predictions are unduly influenced by initial information which is given more weight in the forecasting process	• Start with objective information (e.g. forecasts) • Concentrate on changes and identify the reasons involved
Illusory correlations	Belief that non-existing patterns exist and/or two variables are causally related when they are not	• Verify the statistical significance of patterns/ relationships in terms of changes
Selective perception	People tend to see problems in terms of their own background and experience	• Use people with different backgrounds/ experience who should suggest solutions independently
Regression effects	Persistent increases/decreases might be due to random reasons which, if true, would increase the chance of a future decrease/increase in the opposite direction	• The realization that when errors are random the chances of a negative ones increases when several positive ones have occurred
Attribution of success and failure	Success is attributed to one's skills while failure to bad luck, or someone else's error. This inhibits learning as it does not allow recognition of one's mistakes	• Instead of punishing, encourage people to accept their mistakes and even publicize them so they and others can learn to avoid similar ones in the future
Optimism, wishful thinking	People's preferences for future outcomes affect their forecasts of such outcomes	• Forecasts should be made by a disinterested third party and use more than one person to independently make such forecasts
Underestimating uncertainty	Excessive optimism, illusory correlation, and the need to reduce anxiety result in underestimating future uncertainty	• Estimate uncertainty objectively. Consider various possible futures by considering various scenarios

willing to accept their forecasting biases and initiate steps to avoid them. Another problem is that forecasters, outside the weather field, are afraid to keep track of their predictions because they could be blamed in case they turned out to be wrong. It is necessary to eliminate such fears so that a detailed record of judgmental forecasts is kept and the reasons for possible errors evaluated, not for assigning blames, but for learning and improving future predictions. Being wrong and accepting it is an integral part of the learning process that can be used as feedback to reduce future errors the way it is practiced by weather forecasters.

Averaging and prediction markets. One way of decreasing judgmental biases and reducing the forecasting errors of predictions is by averaging more than one forecaster and/or model [14]. The outcome is not only higher accuracy but also a reduction in the size of forecasting errors, with simple averaging being the best way of combining forecasts [68]. The reason is that averaging cancels out the errors of individuals and/or models and in doing so eliminates the noise from the pattern and improves accuracy.

The value of averaging goes back to Francis Galton in 1906 and has been popularized by Surowiecki's book [87] *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations* as well as by some popular TV programs as those aired by BBC [10]. According to Surowiecki averaging works if

the members of the group are of diverse backgrounds, their decisions are taken independently, and decentralized. Otherwise, the wisdom can turn into madness with catastrophic consequences [57] for those using it. Averaging holds, therefore, both great benefits but also considerable dangers and requires outmost care before it is used by assuring that at least the assumptions of diversity and independence hold.

Prediction markets are forms of averaging the opinions of many people, acting independently, using the web. The idea behind such markets is the creation of a tradable "security" for some event, allowing participants to trade by betting higher or lower, creating a price where supply equals demand, rewarding those who buy low and sell high. Prediction markets are used in addition to presidential and other elections, to Hollywood films Oscar Winners [50], sport events, bitcoins and a variety of other applications as in business firms for forecasting sales, new product introductions or the chances that a project will succeed. In addition, when used in business, prediction markets improve the communication and collaboration among employees and add objectivity to corporate predictions. Employees make their wagers over the Internet betting anonymously by using virtual currency. They are asked to bet on what they believe will really happen, not what they hope will occur or what their boss would like to happen. The payoff for the most accurate player(s) is usually some small prize like a trip or an iPhone. The early results are encouraging as their forecasts seem to be at least as accurate as other alternatives [85].

Prediction markets probably suffer from the same limitations as those mentioned when using the wisdom of crowds, requiring attention when applying their forecasts and being aware of the potential problems when the assumptions of independence and diversity are not satisfied. However, firms can chose participants to assure diversity of backgrounds and expertise while instructing them to avoid common biases. Nevertheless, it seems that prediction markets are promising alternatives to traditional forms of forecasting. A type of such markets is the Good Judgment Project (GJP), initiated by Tetlock et al. [93] which aims to harness the wisdom of the crowd, using the Internet, to forecast political world events.

Tetlock's work encouraged the systematic monitoring of the accuracy of analysts' forecasts and how it would be possible to improve it. This led to the Good Judgment Project (GJP), sponsored by the Intelligence Advanced Research Projects Activities, involving thousands of forecasters making predictions over the Internet on hundreds of questions over time and tracking their accuracy. One of the most interesting findings is that forecasting accuracy does not necessarily improve when analysts have access to highly classified intelligence information as the top forecasters, drawn from universities and elsewhere, performed about 30 percent better than the average obtained by intelligence community analysts who had access to classified data and secret information. But what makes for a good judgmental forecaster in the GJP? Evidence suggests five factors: First, have a highly analytical mind, enjoying thinking through puzzles; Second, possess an active, open mind that applies scientific reasoning based on a rigorous study of the data rather than seeking to accept conventional wisdom; Third, change one's viewpoint as soon as new information becomes available; Fourthly, demonstrate a great deal of curiosity about the world and; Lastly, be able to work effectively in a team of other superforecasters. Although, the GJP is still being evaluated as it is in its fourth year of running it is showing promises as a forecasting tool by (a) aggregating information exploiting the wisdom of the crowd, (b) using the Internet to collect information, (c) encouraging participants to succeed exploiting the concept of commodity markets and (d) identifying and using superforecasters to improve predictive accuracy.

There are similar approaches to Tetlock's GJP aiming at improving forecasting accuracy. One such project is the Geopolitical Forecasting Tournament [75] whose findings indicate that forecasting accuracy and correct calibration of uncertainty can be improved by training, teaming and tracking their performance. Training involves correcting cognitive biases, encouraging forecasters to use reference classes and making them aware of such tools as averaging to exploit the wisdom of crowds. Teaming urges them to share information and discuss the validity of their beliefs. Tracking, finally, identifies the top 2% of performers and putting them together into elite teams to make subsequent predictions. According to the authors of the study such behavioral interventions are advantageous and can be generalized to other forecasting endeavors. However, all of these proposed methods aimed to improve predictive accuracy need to be tested over time by other researchers to prove their superiority to other, reasonable alternatives.

Psychological predictions. There are similarities between political and psychological predictions. Meehl in his now famous little book [73], compared the predictions made by psychologists, based on their obser-

vation and subjectivity of what is wrong with a patient to those grounded on objective, statistical data (called mechanical). He collected and analyzed some twenty studies and discovered that the statistical approach of diagnosis was superior to the traditional, "clinical" approach used by psychologists. Meehl's approach has been replicated by a large number of additional studies. In a meta-analysis Grove et al. [39] summarized the results of 136 studies comparing clinical and statistical predictions across a wide range of situations. They concluded by stating *"We identified no systematic exceptions to the general superiority (or at least material equivalent) of mechanical prediction. It holds in general medicine, in mental health, in personality, and in education and training settings. It holds for medically trained judges and for psychologists. It holds for inexperienced and seasoned judges"* (p. 25).

The findings in the area of psychology are not different than those found in economic/business forecasting, in stock market predictions as well as in political ones, raising questions about the value and contributions of unaided human judgment. Kahneman [44] in his book *Thinking, Fast and Slow* while explaining when we can trust our intuition and when not states *"Statistical algorithms greatly outdo humans in noisy environments for two reasons: they are more likely than human judges to detect weakly valid cues and much more likely to maintain a modest level of accuracy by using such cues consistently"* (p. 241). The fact is that we usually operate in noisy environments making Meehl's conclusions and Grove et al. generalizations valid in the majority of cases to which Kahneman replies with a rule that he encourages us to remember: *"intuition cannot be trusted in the absence of stable regularities in the environment"* (p. 242). Maybe this is something worth remembering when forecasting the future usually done under changing conditions.

4.4. Medical predictions

In medicine there are three types of forecasts required. The first relates to how the doctor makes a diagnosis. The second, how to match such a diagnosis with the correct treatment by providing a recommendation for the most appropriate therapy based on published research findings that are often presented in the form of guidelines written by professional committees of doctors. The big question is how reliable are such findings as new research seems to invalidate old ones. Finally, there is also what is called *preventive* medicine that recommends various procedures and tests to identify potential, future problems before they become serious and take actions to prevent them. This section looks at each of these three types of medical predictions, arguing that their outcome is uncertain as is the case with all other types of forecasts.

How accurate are medical diagnoses? The diagnosis of disease has progressed a great deal with the wide availability of laboratory tests (e.g. blood and urine), equipment like X-rays, ultrasound and MRI machines and PET and CT scanners but significant concerns remain. In a 2013 article Meyer et al. [76], state: *"A total of 118 physicians with broad geographical representation within the United States correctly diagnosed 55.3% of easier and 5.8% of more difficult cases ($P < 0.001$)"* (making an overall average of a 31% success rate). *"Despite a large difference in diagnostic accuracy between easier and more difficult cases, the difference in confidence was relatively small (7.2 vs 6.4 out of 10, for easier and more difficult cases respectively)"* (p. 1952). There are many other diagnostic studies [24] whose conclusions are similar to those of Meyer's et al.

A major finding of the Meyer's et al. study was how little the physicians level of confidence changed from the easy diagnoses to hard ones (7.2 out of 10 for the easy ones and 6.4 out of 10 for the hard ones). This means that with an accuracy rate of only 5.8%, the physicians were still 64% confident that they were right! A low diagnostic accuracy of 5.8% could be tolerated if the physician was for instance, only 10% confident of being right as he would be more likely to order more tests or ask for a second opinion from another doctor. But a 64% confidence would probably exclude such actions thus he would probably proceed with the wrong diagnosis and treatment. Berner and Graber [11] further discuss how overconfidence results in diagnostic errors and what needs to be done to reduce them. It seems that the common problem of overconfidence in forecasting in general is also present in medical diagnosis with its detrimental negative.

How reliable are the findings of medical research? Ioannidis has published widely on the deficiencies of medical research. In his *PLoS Medicine* article [42] he states *"There is increasing concern that in modern research, false findings may be the majority or even the vast majority of published research claims"* (p. 696). In his article in *JAMA* [43], he concludes *"Contradiction and initially stronger effects are not unusual in highly cited research of clinical intervention and their outcomes"* (p. 218). In a 2010 article in the *Atlantic*,

featuring Ioannidis, David Freedman [26] quotes him saying *"that as much as 90 percent of the published medical information that doctors rely on is flawed and that he worries that the field of medical research is so pervasively flawed, and so riddled with conflicts of interest, that it might be chronically resistant to change – or even to publicly admitting that there's a problem"*.

From an epistemological point of view the critical question is how a research finding can be utilized to base therapy when a future, new one could reverse its recommendations? Is there something fundamentally wrong with the practice of medicine that requires a major rethinking on how it is practiced? Should the medical community admit that there is a serious problem facing the profession? In addition, is it possible to ensure that conflicts of interests aimed at increasing the revenues of doctors and pharmaceutical companies will not influence the diagnostic process and the recommended therapy?

Is there some value in preventive medicine? According to the independent Cochrane Foundation *"General health checks involve multiple tests in a person who does not feel ill with the purpose of finding disease early, preventing disease from developing, or providing reassurance… To many people, health checks intuitively make sense, but experience from screening programs for individual diseases have shown that the benefits may be smaller than expected and the harms greater. A possible harm from health checks is the diagnosis and treatment of conditions that were not destined to cause symptoms or death"* ([15], p. 2). For instance yearly checkup examinations started in the early 1920s and have continued since then, although many studies going back to the 1960s have shown no benefits from them. Krogsboll et al. [47] concluded: *"General health checks did not reduce morbidity or mortality, neither overall nor for cardiovascular or cancer causes, although the number of new diagnoses was increased"* (p. 2). Yet, despite evidence against routine annual examinations, many family physicians recommend them [74] exploiting the *"illusion of reassurance"* that a preventive test will catch health problems early, reducing disease and increasing life expectancy. But this has not been the case.

Another frequently recommended preventive test is annual mammography for all women older than 40 (it is estimated that 85% of women in the USA each year are screened for breast cancer). Lately, the starting age has been raised to 50 and the interval of the screening to two years instead of annually. But there is still an intense argument for the value of any screening for women at any age with those opposing it saying that there is strong evidence that the potential harm from screening is considerably greater than the benefits. Gøtzsche in his book [35] *Mammography Screening: Truth, Lies and Controversy* states:

> *"If we wish to reduce the incidence of breast cancer, there is nothing as effective as avoiding getting mammograms. It reduces the risk of getting breast cancer by one-third"* (p. 349).

Prostate cancer tests for men are also common as it is estimated that close to 52% of men are tested annually at a cost exceeding half a billion for Medicare patients alone in the USA. Richard Ablin, the inventor of the PSA test being used to diagnose prostate cancer, in his book [1] *The Great Prostate Hoax* states:

> *"The ability of the PSA test to identify men with prostate cancer is slightly better than that of flipping a coin. And its continued use as a routine screening tool is nothing short of a national health disaster"* (p. 6).

Albin's conclusion is that the screening is done because of financial interests as it increases the number of additional tests and prostatectomies. He asserts that *"Without radical prostatectomies, more than half of all the urology practice in the United States would go belly-up"* (p. 42).

Preventive testing is another medical area requiring fundamental rethinking with the advice of epidemiologist McPherson being that *"reducing incidence (diagnosis of cancers) must be the primary goal, with reducing mortality an important but secondary end point"* (pp. 233–235) (see [35], p. 361). The question is if vested interests from both doctors and pharmaceutical firms would allow any changes in established preventive medicine practices?

5. Assessing uncertainty and preparing to face risk

> "You can never be certain about uncertainty."

So far in this paper we have referred to "uncertainty" a great number of times and although its every day meaning may be understood this is not the case with its scientific and technical one which differs among the various fields of endeavor. Uncertainty lacks a unique definition and clear meaning that varies when used in different fields such as philosophy, physics, information science, economics, statistics, finance, poli-

tics, psychology, engineering and insurance. But even within the same field, for example physics, uncertainty can have completely different meanings as in Newtonian mechanics (no uncertainty), quantum theory (uncertainty principle), or in Prigogine's views that uncertainty is pervasive and deeply embedded within the core of reality and as such plays a critical role in such diverse areas as creativity and progress.

The everyday definition of uncertainty and that of different fields are presented below:

- *In common sense*: The lack of certainty; Inability to exactly describe present states, or sufficiently predict future ones; Statements that we are not sure to be true or false.
- *In Physics*
  - ∗ *Newtonian* (*Classical*) *Mechanics*: Its laws account for all the motions of the celestial bodies, and of our sea in a deterministic manner excluding uncertainty completely.
  - ∗ *Heisenberg's Uncertainty Principle*: We cannot know with precision both the location and momentum (speed and direction) of a subatomic particle, pointing to a fundamentally unknowable uncertainty.
  - ∗ *Prigogine End of Certainty*: Determinism is not a viable scientific belief. We do not live in a predestined world or in one of pure chance. Uncertainty creates novelty through an unplanned, unguided, creative power embedded in the universe itself.
- *In Information Science*: Claude Shannon defined "entropy" as a measure of uncertainty with respect to some variable or event, the greater the uncertainty, the greater the entropy.
- *In Statistics and Finance*: The variability around the mean as measured by the variance or the standard deviation.
- *In forecasting*: The confidence intervals, with a certain degree of conviction, around the most likely predicted value.
- *In Meteorology*: A single probability of the likelihood of occurrence of the predicted event.
- *In Decision Theory*: A set of possible states or outcomes with corresponding probabilities assigned to each of them.
- *In Judgmental Psychology*: The unknown influence of biases and irrationality in decision making.
- *In economics*: The "Knightian uncertainty" is defined as a non-measurable risk.
- *In Philosophy*: Part of the uncertainty in philosophy derives from the very nature of the questions that it undertakes to answer (e.g., has the universe any unity of plan or purpose? is consciousness a permanent part of the universe? are good and evil of importance to the universe or only to man?).

Uncertainty relates also to risk.

- *Risk*: The outcome from assessing uncertainty to figure out its implications. Some believe that risk should not be presented unless it is measurable. However, its magnitude can be appreciated even if it cannot be quantified. Furthermore, some only refer to negative risks while there can also be positive one, in the form of opportunities, when dealing with uncertainty (e.g., when starting a new firm, or from buying stocks whose values can appreciate).

Uncertainty relates to reducing risk and preparing to cope with it.

- *Insurance*: The price to be paid to reduce physical or other risks.
- *Hedging*: The price to be paid to reduce investment risks.
- *Weather Derivatives*: Buying contracts that pay out money according to stipulated weather conditions.

Uncertainty and Extreme Value Theory (EVT). In many cases using the average or median uncertainty to assess risk is of little or no value as in the catastrophic flooding of New Orleans on August 29, 2005, caused by hurricane Katrina's extreme, nine meter waves that overwhelmed the city's levees and drainage canals. It was the extreme height of the waves and the equally heavy rain responsible for breaking the levees that had withstood many other hurricanes for more than 36 years since the last New Orleans' flooding. This is the contribution of EVT, a branch of statistics dealing with extreme deviations (see [84]) rather the average or the median as it concentrates on the probability of the extreme events in the historical data. ETV is used in disciplines such as floods and hurricane predictions [28], finance and insurance [19], earth sciences, and geological engineering and finance [86]. EVA has been used extensively in the field of hydrology [32] to estimate the probability of unusually large flooding events, such as the 100-year flood as well as estimating rainfall extremes [45]. Extreme Value Theory is of particular importance for the realistic assessment of uncertainty as it is being influenced by extreme values in a non-linear

fashion. Moreover, the possibility of non-stationarity in past patterns needs to be added as it may further increase extreme values. Extreme values amplify tail risks, making their estimation critical. EVT becomes necessary for both natural phenomena and events in the social sphere in order to come up with realistic assessments of uncertainty and risk.

Uncertainty and thin/fat tailed errors. There are two types of forecasting errors with critical implications on the uncertainty and risk of both natural and social events. Thin tailed errors characterize the great majority of situations and fluctuate within the statistical limits of the normal, or other distributions. Their variance is finite and measurable, allowing the estimation of the size of forecasting errors, or the construction confidence intervals around the most likely forecasts. Thin tailed errors are typical of the "usual", or "normal" conditions that prevail most of the time in the natural environment and in the economy and business. The great plethora of statistical models employed in economics, business, finance and other fields assume thin tailed errors and normal distributions restricting their applicability and realism to usual/normal circumstances. Fat tailed errors, on the other hand, prevail during periods of environmental disturbances (heavy storms and hurricanes, big floods, earthquakes etc.) or during economic and other crises (or booms). Their variance is infinite, excluding the estimation of the forecasting errors or the construction of confidence intervals. Sometimes, the size of such errors exceeds the six or seven standard deviations, indicating that their chance of occurrence is less than one in a billion. Clearly, uncertainty and risk are huge in the case of fat tails and cannot be ignored as is usually done by treating them as outliers. Considerable efforts must, therefore, be made to accept the existence of fat tails and develop models that take into account their huge fluctuations (see for example [91]).

Given the plethora of definitions of uncertainty and types of forecasting errors, it is the purpose of this section to clarify "uncertainty", relate it to forecasting, illustrate the risks (both negative and positive) involved, demonstrate their known consequences, imagine unknown ones and consider what could be done to prepare to face both the known and unknown risks, accepting that uncertainty is not static but changing over time.

5.1. Natural events: Uncertainty, risk and preparation

Although there are differences, the natural events considered above cannot be predicted at all, or to a limited degree, with the only exception being that of immediate and short term weather under normal conditions. Of course, we can make estimates from historical records where earthquakes are most likely to strike (according to the USGS, more than 50% of magnitude 3.5 and greater earthquakes hit Alaska while none occurred in many states, including Wisconsin, Vermont and North Dakota). However, no accurate predictions can be made as to the exact timing, location and magnitude of earthquakes in Alaska or anywhere else. At the same time there is little to no chance of earthquakes hitting, say, Wisconsin. Practically the same situation holds for most natural disasters as the chance of avalanches is zero anywhere apart from mountainous, high altitude areas which are, however, unlikely to be flooded. The critical question becomes, therefore, how to evaluate uncertainty realistically, properly assess the risk involved and take actions to reduce the negative consequences in the areas affected by natural disasters. The uncertainty, risk and possible actions to be prepared vary with each type of potential calamity.

No prediction whatsoever (earthquakes) but preparation. As mentioned the location, timing and magnitude of earthquakes cannot be predicted, creating huge uncertainty in particular for strong ones. The risk for deaths, property and infrastructure is usually estimated from historical records of past ones which, however, do not take into account that a stronger one than those in the past can occur. Preparing for future earthquakes can take two forms. The first is in the form of building codes so that buildings and infrastructure can withstand earthquakes of at least certain magnitudes while minimizing the chance of fires that usually follow earthquakes. The second is in the form of drills aimed at preparing people to face an actual earthquake. With technological advancements and the higher construction expenditures for buildings and infrastructure we can withstand stronger earthquakes and reduce deaths.

No prediction some limited warning (tornados). Tornados are Black Swans as their specific location, timing and path cannot be predicted, creating vast uncertainty while the risk of potential destruction can be considerable even though it is limited to small areas. In some cases an alert can be issued to warn of forthcoming tornados, but only once funnel clouds have been spotted or seen on radars. Apart from communicating such information with sirens, if such exist, or announcing it over radios and TVs, not much can be done apart from the advice given to take immediate precautions as the time between the appearance of funnel clouds

and the actual tornado is slight. However, the area of the warning of possible tornados has been reduced to a county level improving significantly its value [16]. The time of the warnings has also improved being on average thirteen minutes, versus five minutes in the past. The major difficulty is still the number of false alarms that discourage people to follow the issued warnings.

No prediction but monitoring (tsunamis, volcanos and floods). Monitoring can be used to reduce uncertainty and risk in the case of tsunamis. Warnings are issued after a strong earthquake has hit a certain area while also using buoys to monitor for possible big waves. Such warnings urge people to move to high grounds to avoid the overflowed coastal areas. But as with tornados there are plenty of false alarms discouraging people from following the warnings. Volcano lava is easier to monitor so that an evacuation order can be issued for affected areas. The same is true for floods although sometimes they do not leave enough time for evacuation warnings to be implemented in an orderly fashion.

No prediction but active observation (forest fires, avalanches). Through continuous observation forest fires can be caught early and extinguished while avalanches can be prevented by explosives provided there are no people in the affected area. Such measures reduce uncertainty and avoid or minimize losses.

No prediction followed by effective monitoring and some warning (hurricanes). The timing of hurricanes cannot be predicted. However, once a satellite indicates that a certain hurricane is developing there is a close watch of its progress and once confirmed a continuous monitoring of its strength and path starts. This is done by weather satellites, specially equipped planes, dropsondes (GPS-enable sensors on parachutes emitted at the center of the hurricane to provide accurate measurements among others, of its strength and direction) and drones. The outcome of such close tracking coupled with the usage of more powerful computers and better models has increased the forecasting of hurricanes from three to five days in advance while it has also narrowed down the range they will hit from 480 kilometers to 400 kilometers. As new specialized weather satellites, land stations, more powerful computers as well as more sophisticated weather models are planned, the forecasting of hurricanes will be extended to seven days while also narrowing their hitting range [58].

Once underway some prediction coupled with preventive action (droughts and floods). Extended droughts can affect large areas and for long periods of time causing considerable hardship and disputes between farmers, city-dwellers and environmentalists on how water should be consumed. Once droughts have been confirmed and the risks of running out of water assessed, the emphasis shifts to the adoption of the most appropriate water management measures to minimize the negative consequences. In addition, there are long term drought models predicting that the American Southwest and central Great Plains will experience extensive droughts in the second half of this century [25], without saying anything about what could be done to deal with them, or mentioning the uncertainty of the forecasts. At the same time, there has been useful work on how to prevent floods using the Extreme Value Theory (EVT) to estimate upper flood bounds and take corrective actions before it is too late to avoid floods.

5.2. Social events: Uncertainty, risk and preparation

Social events differ from natural ones by the fact that people can influence their future course, adding an additional level of difficulty in both predicting them and assessing uncertainty. According to Table 1 it is clear that uncertainty and risk vary significantly in each of its four quadrants. Most importantly, however, uncertainty and risk are not constant over time as the situation under consideration can move from one quadrant to another. The Appendix describes the example of Lehman Brothers that illustrates such movement among the four quadrants and the consequences of uncertainty and risk.

Cost–benefits of preparation. Accepting uncertainty and evaluating the risks involved is of little value unless concrete steps are taken to prepare to mitigate such risks. The problem is that in the great majority of cases to do so involves a cost that must be compared to the perceived benefits. For instance, the costs of different fire insurance policies can be compared to their expected benefits, assuming various levels of destruction, and the ability of the homeowner to pay. The same is true with building codes that increase construction costs as a function of the magnitude of the expected earthquakes, or additional cost for increasing the height of river banks to reduce the risk of floods. According to Casti [13] hidden catastrophes loom and can lead to the collapse of everything. In such a case, there is little that can be done on a practical level to

prepare to face such events. On the contrary, Taleb [90] provides advice of what can be done to assess uncertainty and be prepared to face risks by adopting antifragile strategies. This is particularly true in the business and economics areas where uncertainty, risk and preparation are often ignored, believing that bad things cannot happen to us (see [67]). In final analysis balancing the expected costs and benefits is of critical importance and must become a priority in preparing for the great range of eventual risks confronting policy and decision makers.

5.3. Conclusions and directions for future research

This paper surveyed forecasting in a good number of diverse fields in both natural and social sciences and discussed their accuracy and reliability. A summary of the major findings of the survey and the challenges facing the field is presented below:

- With the exception of weather forecasting, there is little or no progress in improving the accuracy of predictions over time. Studies going back to the late 1970s [7,62] show large forecasting errors, the inability of complex and statistically sophisticated methods to outperform simple ones, including naïve benchmarks, as well as little value from the numerous public forecasts available to policy and decision makers [7]. Such conclusions are still valid after many decades of research and practical usage (see special issue of the *Journal of Business Research* [38]). In addition academic research showing how to improve forecasting accuracy has been largely ignored by researchers and practitioners.
- Judgmental forecasts, influenced by human biases and limitations, are no more accurate than statistical ones.
- Combining forecasts seems to improve accuracy as long as they are independent and collected from a diverse group. This is particularly true in judgmental predictions where the wisdom of the crowds (i.e. the averaging of many predictions) usually provides more accurate forecasts than the best individual, while also reducing the variance of forecasting errors. The problem is, however, that sometimes such wisdom turns into madness when crowds follow herd instinct, what Keynes has called animal spirit, and overreact to both good and bad news.
- Uncertainty is seriously underestimated by both statistical models as well as judgmental forecasters with the exception again of weather ones.
- There are two distinct types of forecasting situations requiring different approaches and models. The first refers to predicting normal conditions with established patterns and existing relationships in a stable, steady situation. The second is during unusual situations characterized by transient, changing patterns/relationships. In the first category we have normal weather and typical economic conditions while the second can include storms and hurricanes, or recessions and other crises (or booms).
- In economy and business recessions and crises (or booms) cannot, therefore, be treated as outliers but must be predicted using a different approach and model, a point made forcefully by Buchanan [12].
- The accuracy and uncertainty in predictions varies considerably depending on the time horizon of predictions (immediate, short, medium and long term) presenting different problems and challenges to forecasters. Furthermore, such accuracy/uncertainty varies from one field to another (e.g., the long term consistency of stock and commodity markets versus the immediate term in weather predictions).
- Weather forecasting provides the exceptional case of improvements towards more accurate and better calibrated predictions, presenting an example to be followed by other disciplines. According to Silver [81] and others the reasons for the superior performance of weather forecasters are that they:
  - Accept the complexity and therefore the limits of predictability.
  - Start from the present and base their forecasts on current meteorological conditions.
  - Compare forecasting accuracy to simple benchmarks to determine improvements.
  - Evaluate the forecasts versus actuals, using continuous feedback, and attributing accuracy improvements to models and/or human judgment.
  - Make sure the uncertainty in their forecasts is unbiased adjusting their calibration on a continuous basis using feedback.
  - Use a different approach and model for normal weather predictions versus unusual ones as those of storms and hurricanes.

- The forecasting errors during "normal" conditions can be characterized as thin tailed while those of unusual one behave in a completely different manner following fat tails.
- Researchers in hydrology and other disciplines have been successfully using Extreme Value Theory (EVT) to estimate uncertainty and come up with realistic assessment of risks that accept fat tail errors and avoid the trap of "average" ones that underestimate uncertainty and risk significantly. Their findings have great potential and can be applied to other forecasting areas.

The future of forecasting: The challenge for practitioners and researchers. Will the accuracy and reliability of predictions be improved beyond weather ones? In this section we will concentrate on forecasts in the economic and business areas, being the most familiar to us, by analyzing major roadblocks and proposing possible improvements. We will distinguish our recommendations to researchers and practitioners.

The major objective of researchers in forecasting is to advance their career by publishing in high impact referred journals. This moves them towards theoretical research and academic publications that provide little practical value. A common mistake is sophisticated models that fit past data well but cannot predict the future accurately as it is well known that better model fit has little or no relation to more accurate, post sample predictions. Authors must judge their models for their real forecasting performance rather than how well they explain the past, by minimizing mean squared errors or maximizing the value of R^2, while journal editors must not accept papers that do not provide out of sample comparisons. Moreover, ranges of uncertainty must also be included around those of point forecasts, stressing their limitations and the fact that uncertainty abounds in any and all future predictions. In addition, every effort must be made to avoid underestimating uncertainty even though there may be considerable pressure for doing so.

One research conclusion that has been widely accepted is the one reached in the M-Competitions stating that simple methods are at least as accurate as sophisticated ones. The 2015 special issue of the *Journal of Business Research* on "Simple Versus Complex Forecasting" reaffirms that conclusion and the inability of complex models to improve on the post sample accuracy of predictions. The challenge is to go beyond simple models to further improve such accuracy by studying and learning as much as possible from weather forecasters who have been improving over time the accuracy of their predictions. In addition the emphasis must shift from accuracy to the realistic estimation of uncertainty. This would require a continuous, honest evaluation of the forecasts being made in the economic and business fields and using feedback to improve not only accuracy, but most importantly the better calibration of uncertainty. In this direction practitioners can contribute by keeping track of their forecasts distinguishing them into the three categories of statistical, judgmental and final, after having incorporated judgmental inputs to the statistical predictions. It is absolutely necessary to do so in order to assess where improvements come from in order to avoid biases and better calibrate uncertainty. Unfortunately, forecasters are not comfortable with keeping track of their forecasts fearing that they will be accused in case of errors. But no progress can be achieved unless such fear is overcome and an objective evaluation of forecasts and uncertainty is made on a continuous basis. Forecasting errors cannot be improved unless there is continuous learning from past mistakes. Finally, business and economic forecasters can learn from hydrology and apply EVT to assessing uncertainty and risk more realistically.

Advances also need to be made in the way that forecasts and uncertainty are communicated to the end users. Weather forecasters have been concerned for instance with how to present the trajectory of forthcoming hurricanes not as a single line but rather as the possible areas that may be hit by the hurricane (called the uncertainty cone). Thus warning the residences of the possible areas to be affected and allowing them to prepare themselves to face the hurricane or evacuate to a safer place. The problem comes in cases of false alarms that require the evacuation of people when no hurricane hits and how they respond the next time that a hurricane alarm is issued. It is therefore important to balance the risk of the actual event and the possibility of a false alarm as there are costs and benefits involved.

Finally, the greatest of all challenges facing the field of economic and business forecasters is to accept that recessions/crises and booms are not outliers but part of their task, and to develop appropriate, even crude models to help them in their predictions. The possibility of the effective monitoring of, say, recessions as is practiced once a hurricane has been confirmed can be useful to at least know where the economy is headed. Similarly, the ideas of Extreme Value Theory may be useful in estimating the extent of the forthcoming recession. In addition the usual forecasting tasks must be expanded by the continuous monitoring of the econ-

omy, industry and business, the constant assessment of uncertainty and the realistic evaluation of risk to be prepared for action when it becomes necessary. Methods and practices from the forecasting of natural events can in our view contribute positively to improving the fields of economic and business forecasting.

Acknowledgements

The authors would like to thank Professors Kesten Green, Robert Fildes, Ari Syntetos, Fotios Petropoulos and anonymous referees for their helpful comments and insightful suggestions for improving this paper.

Appendix. Lehman Brothers' trajectory from the Known/Knowns to the Unknown/Unknowns quadrant

Lehman Brothers started in Montgomery, Alabama in 1844 as a general store, evolved into a commodity broker and then a firm trading securities as it became a member of the New York Stock Exchange in 1887. Since then it expanded into financing major railroad and oil projects and eventually became a major global bank that celebrated its 150 years of history in 2000. Could the uncertainty and risk associated with Lehman Brothers have been assessed objectively by third parties or the company itself that according to its brochure claimed *"The effective management of risk is one of the core strengths that has made Lehman Brothers so successful"?*

In terms of Table 1 Lehman Brothers operated in the known/knowns quadrant for over 150 years. Variables like its daily share price, monthly revenues, quarterly dividends or its yearly profits were estimated more or less accurately and the uncertainty surrounding such estimates were reasonable and within acceptable limits. The risk associated with the uncertainty facing the firm were assessed over time with no reason for alarm as the firm was growing satisfactorily in both revenues and profits. Lehman's reported record profits in 2005, 2006 and 2007 at which year they exceeded $4 billion on net revenues of close to $20 billion.

But the normal settings under which the firm operated for over one and a half centuries terminated with the subprime crisis that hit the US economy and the ensuing 2007–2008 financial crisis. The firm then moved from the known/knowns quadrant to the unfamiliar known/unknowns one as the economy entered the steep recession, the uncertainty and risks associated with its operations multiplied and the implications of the crises could not be assessed accurately. Finally, as the crisis deepened further, Lehman Brothers were pushed into the unknown/unknowns category that eventually forced it into bankruptcy, a Black Swan nobody could have predicted would happen to such a blue chip firm. Its stock price dropped from more than $86 in February 2007 to practically zero a year and a half later in September 15, 2008.

What is clear from the Lehman Brothers example is that uncertainty and risk cannot be assumed to be constant. Instead they vary significantly depending on changes in economic, political, operational and other conditions. They must, therefore, be assessed differently depending on the specific, prevailing situation. Doing otherwise would be as equally futile as trying to assess the uncertainty, risk and damages of a hurricane assuming normal weather conditions. Next the uncertainty and risk of Lehman Brothers in each of the four quadrants are evaluated to demonstrate that uncertainty is a highly complex issue that defies conventional logic as it is being transformed significantly with the changing conditions that affect it and worse doing so in little time. The same is true with the assessment of risk that also varies widely ending up being infinite before an eventual bankruptcy.

The Known/Knowns. For the over 150 years Lehman operated in this quadrant, there was some uncertainty and share prices that sometimes exceeded but no major problems. On the contrary, the firm was showing considerable opportunities and substantial growth prospects as its revenues and profits were growing handsomely, a fact reflected in its share price that skyrocketed and went from less than $5 in 1994 to over $86 in February 2007. In addition, its dividends more than doubled from $0.24 in 2003 to $0.60 in 2007, four years later, reflecting the healthy financial position of the firm. There were no signs of trouble according to its 2007 Annual Report that stated: *"In 2007, Lehman Brothers produced another year of record net revenues, net income, and earnings per share and successfully managed through the difficult market environment"* later continuing: *"We effectively managed our risk, balance sheet, and expenses. Ultimately, our performance in 2007 was about our 'One Firm' sense of shared responsibility and careful management of our liquidity, capital commitments, and balance sheet positions. We benefited from our senior level focus on risk management and, more importantly, from a culture of risk management at every level of the Firm"*. By the time, however, the 2007 Annual Report was distributed

Lehman Brothers risk management was in serious trouble as the real estate bubble in the USA had started to burst.

There are no guarantees how long a firm like Lehman will stay in the known/knowns category, just as no one can be certain that a storm or hurricane will not develop after a long period of normal weather. It is a fact that the majority of organizations operate most of the time under normal conditions during which uncertainty can be adequately assessed, and risk adequately evaluated. It is the same with normal weather that prevails most of the time and although a switch from normal to stormy is to be expected, its timing cannot be known, creating an additional, significant level of uncertainty and risk that has been ignored until now in both the academic literature and among business gurus and writers. But how does a firm move from the known/knowns category to others? In other words what are the factors that brought Lehman Brothers to bankruptcy, at least in hindsight?

The Known/Unknowns. Between 1997 and 2006 US home prices increased by 124%, partly due to the government's encouragement to increase home ownership. For achieving such an objective banks and mortgage institutions were offering subprime loans to people who could have difficulty repaying them and who were charged higher interest rates in order to compensate for the higher credit risk. Consequently, such subprime loans were packaged into mortgage-backed securities and sold to third parties, often as high grade investments. As long as home prices were climbing everybody was happy seeing their property value increasing. However, the size of these subprime loans was growing dangerously and became a major problem as a rising number of subprime borrowers were unable to meet their mortgage payments and were defaulting. This became the contributing factor that lead to the financial crisis of 2007–2008 and the ensuing Great Recession. Subprime lending started to seriously affect Lehman Brothers that was heavily exposed to property derivatives, creating a new situation for the firm that moved it to the known/unknowns category.

Lehman's financial strength started to deteriorate as it was reflected by its share price which fell from its height of $86.2 in February 2007 to below $70 a few months later and then started fluctuating around $60 for the remaining part of the year. Things got worse with the arrival of 2008 as its share price started a severe, continuous decline reaching $12 in the middle of July 2008 and eventually reaching zero in the middle of September with its bankruptcy. The journey from the high of $86.2 in February 2007 to zero in September 2008 lasted less than nineteen months and caught everyone by surprise, including Lehman's CEO Dick Fuld who blamed everyone but himself for his firm's bankruptcy [27].

The Unknown/Knowns. In judgmental psychology it is known that people tend to undervalue dangers such as disease, serious accidents, recessions and other hardships whose existence they accept but believe cannot happen to themselves [18,67]. This relates to the illusion of control which is defined as the "expectancy of a personal success probability higher than the objective probability would warrant" ([48], p. 313) that encourages people to believe that things are under "control" and also believing the "it cannot happen to me" syndrome. They are not willing, therefore, to accept available evidence clearly showing that people cannot avoid "bad things" happening to themselves. Lehman's CEO, Dick Fuld, in a recent talk declared that there were no risk management problems with his firm as "every one of Lehman's 27,000 employees was in risk management". The firm ignored messages from the millions of investors who forced the price of its stock to drop to less than half its value in just a year, from February 2007 to February 2008. Its top management did nothing to reduce its high leverage that was increased instead from 26.2 in 2006 to 31 before its bankruptcy. The fact that high leverage rates create added risk was ignored by a firm priding itself for its expertise in risk management while it also overlooked the fact that during periods of financial downturns extra liquidity and added capital may be needed. These knowns were ignored by Lehman Brothers although they should not have. This is what the Unknown/Knowns category stands for as people are unwilling to accept known facts that unconsciously avoid to take into account and which move them into the Unknown/Knowns category. This is what happened to Lehman's as top management that was not willing to accept the evidence (the knowns), preferred to believe that "bad things, although possible to others, cannot happen to their firm" [72]. The end result was that Lehman moved into the unknown/unknowns quadrant which lead to its eventual bankruptcy, mainly because its top management could not accept the increasing uncertainty and risk facing their firm.

The Unknown/Unknowns. Lehman's bankruptcy caught everyone by surprise, including its CEO, as he believed that the company was "too big to fail" and that the FED, fearing catastrophic financial consequences,

would come to its rescue. Betting on such a belief, its CEO Dick Fuld rejected an acquisition offer of $18 per share from the Korean Development Bank in August of 2008 as too low while looking for interim solutions that failed to save the firm. The aftermath was the near collapse of the global financial system that was only saved through the massive insertion of liquidity by the FED into the US financial system. Lehman moved from the known/knowns to the unknown/unknowns in a little over eighteen months, raising fundamental questions about the "assessment of uncertainty" as well as the evaluation of risk as everything can change in even short periods of time. Critical to this transition was the "unknown/knowns" as its top management refused to consider simple known facts that govern periods of recessions and firms in dire financial positions.

References

[1] R.J. Ablin, *The Great Prostate Hoax: How Big Medicine Hijacked the PSA Test and Caused a Public Health Disaster*, Palgrave Macmillan, New York, 2014.

[2] P.G. Allen and B.J. Morzuch, Twenty-five years of progress, problems, and conflicting evidence in econometric forecasting. What about the next 25 years?, *International Journal of Forecasting* **22**(3) (2006), 475–492.

[3] Aristotle, Meteorology, 350 BC, available at: http://classics.mit.edu/Aristotle/meteorology.1.i.html. (Translated by E.W. Webster.)

[4] J.S. Armstrong, Forecasting with econometric methods: Folklore versus fact, *Journal of Business* **51**(4) (1978), 549–564.

[5] J.S. Armstrong and R. Fildes, Making progress in forecasting, *International Journal of Forecasting* **22**(3) (2006), 433–441.

[6] J.S. Armstrong, K.C. Green and A. Graefe, Golden rule of forecasting: Be conservative, *Journal of Business Research* **68**(8) (2015), 1717–1731.

[7] W. Ascher, *Forecasting: An Appraisal for Policymakers and Planners*, Johns Hopkins Univ. Press, Baltimore, 1978.

[8] R. Batra, *The Great Depression of 1990*, Simon & Schuster, New York, 1988.

[9] J. Battelle, *The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture*, Penguin Group, New York, 2005.

[10] BBC, The code – The wisdom of the crowd, 2011, available at: https://www.youtube.com/watch?v=iOucwX7Z1HU.

[11] E.S. Berner and M.L. Graber, Overconfidence as a cause of diagnostic error in medicine, *The American Journal of Medicine* **121**(5A) (2008), S2–S23.

[12] M. Buchanan, *Forecast: What Physics, Meteorology, and Natural Sciences Can Teach Us About Economics*, Bloomsbury, New York, 2013.

[13] J. Casti, *X-Events: The Collapse of Everything*, Harper Collins, New York, 2012.

[14] R.T. Clemen, Combining forecasts: A review and annotated bibliography, *International Journal of Forecasting* **5** (1989), 559–583.

[15] Cochrane Library, General health checks in adults for reducing morbidity and mortality from disease, 2015, available at: http://www.cochranelibrary.com/editorial/10.1002/14651858.ED000047.

[16] T.A. Coleman, K.R. Knupp, J. Spann, J.B. Elliott and B.E. Peters, *The History (and Future) of Tornado Warning Dissemination in the United States*, American Meteorological Society, 2011.

[17] E. Dimson, P. Marsh and M. Staunton, *Triumph of the Optimists: 101 Years of Global Investment Returns*, Princeton Univ. Press, Princeton, NJ, 2002.

[18] D. Dolinksi, W. Gromski and E. Zawisza, Unrealistic pessimism, *Journal of Social Psychology* **127** (1987), 511–516. doi:10.1080/00224545.1987.9713735.

[19] P. Embrechts, C. Klüppelberg and T. Mikosch, *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin, 1997.

[20] EPA, Environmental Protection Agency, Climate change is happening, 2015, available at: http://www3.epa.gov/climatechange/basics/.

[21] E.F. Fama, Efficient capital markets: A review of theory and empirical work, *Journal of Finance* **25** (1970), 28–30. doi:10.2307/2325486.

[22] R. Fildes and P. Goodwin, Forecasting: Robert Fildes and Paul Goodwin investigate the effectiveness of firms' judgmental adjustments to computer-generated demand forecasts, *Financial Management* (2007).

[23] R. Fildes and S. Makridakis, The impact of empirical accuracy studies on time series analysis and forecasting, *International Statistical Review* **63**(3) (1995), 289–308.

[24] W. Fink, V. Lipatov and M. Konitzer, Diagnoses by general practitioners: Accuracy and reliability, *International Journal of Forecasting* **25** (2009), 784–793.

[25] M. Fischetti, U.S. Droughts Will Be the Worst in 1,000 Years: The Southwest and central Great Plains will dry out even more than previously thought, *Scientific American* (2015), available at: http://www.scientificamerican.com/article/u-s-droughts-will-be-the-worst-in-1-000-years1/.

[26] N.D. Freedman, *The Atlantic* (2010), available at: http://www.theatlantic.com/magazine/archive/2010/11/lies-damned-lies-and-medical-science/308269/.

[27] S. Gandel, Dick Fuld: 'Lehman was not a bankrupt company', *Fortune* (2015), available at: http://fortune.com/2015/05/28/dick-fuld-lehman-brothers/.

[28] A.R. Ganguly, E.A. Kodra, A. Agrawal et al., Toward enhanced understanding and projections of climate extremes using physics-guided data mining techniques, *Nonlinear Processes in Geophysics* **21** (2014), 777–795.

[29] R. Geller et al., Earthquakes cannot be predicted, *Science Online* **275**(5306) (1997), 1616, available at: http://scec.ess.ucla.edu/~ykagan/perspective.html.

[30] G. Gigerenzer, *Risk Savvy: How to Make Good Decisions*, Penguin Group, New York, 2014.

[31] J.K. Glassman and K.A. Hassett, *Dow 36,000: The New Strategy for Profiting from the Coming Rise in the Stock Market*, Random House, New York, 1999.

[32] J. Goldstein, M. Mirza, D. Etkin and J. Milton, Hydrologic assessment: Application of extreme value theory for climate extremes scenarios construction, in: *14th Symposium on Global Change and Climate Variations, 83rd Annual Meeting of the American Meteorological Society*, 2003.

[33] Google Images (retrieved 1 September 2015), available at: https://www.google.com.cy/search?q=ice+core+cycles&biw=1920&bih=955&tbm=isch&tbo=u&source=univ&sa=X&ved=0CDYQsARqFQoTCLeinuv-9ccCFUi2Ggod9IkGYA.

[34] J.G. Gooijer and R.J. Hyndman, Abstract 25 years of time series forecasting, *International Journal of Forecasting* **22**(3) (2006), 443–473.

[35] P.C. Gøtzsche, *Mammography Screening: Truth, Lies and Controversy*, Radcliffe Medical Press, London, 2012.

[36] J.R. Graham and C.R. Harvey, Market timing ability and volatility implied in investment newsletters' asset allocation recommendations, *National Bureau of Economic Research* (1997), Working Paper 4890.

[37] K.C. Green and J.S. Armstrong, Forecasting global climate change, in: *Climate Change: The Facts 2014*, A. Moran, ed., IPA, Melbourne, Australia, 2014, pp. 170–186.

[38] K.C. Green and J.S. Armstrong, Special Issue on Simple Versus Complex Forecasting, *Journal of Business Research* **68**(8) (2015), 1657–1818.

[39] W.M. Grove et al., Clinical versus mechanical prediction: A meta-analysis, *Psychological Assessment* **12**(1) (2000), 19–30.

[40] D.J. Hand, *The Improbability Principle: Why Coincidences, Miracles, and Rare Events Happen Every Day*, Farrar, Straus and Grioux, New York, 2014.

[41] C. Hsiao, Open forum on the current state and future challenges of economics, *Journal of Econometrics* **100**(1) (2001), 1–112.

[42] J.P. Ioannidis, Why most published research findings are false, *PLoS Medicine* **2** (2005), e124.

[43] J.P. Ioannidis, Contradicted and initially stronger effects in highly cited clinical research, *JAMA Internal Medicine* **294** (2005), 218–228.

[44] D. Kahneman, *Thinking, Fast and Slow*, Farrar, Straus and Giroux, New York, 2013.

[45] A. Kamali, Review of progress in analysis of extreme precipitation using satellite data, *Open Transactions on Geosciences* **1**(2) (2014), 23–33.

[46] R. Kosowski et al., Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis, *The Journal of Finance* **61**(6) (2006), 2551–2595.

[47] L.T. Krogsbøll, K.J. Jørgensen, L.C. Grønhøj and P.C. Gøtzsche, General health checks in adults for reducing morbidity and mortality from disease, *Cochrane Database of Systematic Reviews* **2012**(10) (2012), CD009009. doi:10.1002/14651858.CD009009.pub2.

[48] E.J. Langer, The illusion of control, *Journal of Personality and Social Psychology* **32** (1975), 311–328. doi:10.1037/0022-3514.32.2.311.

[49] M. Lawrence, M. O'Connor and B. Edmundso, A field study of sales forecasting accuracy and processes, *European Journal of Operational Research* **122**(1) (2000), 151–160.

[50] D. Leonhardt, Oscars 2015: An excellent night for prediction markets, *The New York Times* (2015).

[51] M. Lewis, *Flash Boys: A Wall Street Revolt*, 1st edn, Norton and Company, New York, 2014.

[52] E. Lorenz, Deterministic nonperiodic flow, *J. Atmos. Sci.* **20** (1963), 130–141.

[53] P. Loungani, The arcane art of predicting recessions, *Financial Times* (2000).

[54] R. Lucas, Econometric policy evaluation: A critique, in: *The Phillips Curve and Labor Markets*, K. Brunner and A. Meltzer, eds, Carnegie–Rochester Conference Series on Public Policy, Vol. 1, American Elsevier, New York, 1976, pp. 19–46.

[55] E.J. Lusk and J.S. Neves, A comparative ARIMA analysis of the 111 series of the Makridakis competition, *Journal of Forecasting* **3**(3) (1984), 329–332.

[56] P. Lynch, *The Emergence of Numerical Weather Prediction: Richardson's Dream*, Cambridge Univ. Press, London, 2014.

[57] C. MacKay, *Extraordinary Popular Delusions and the Madness of Crowds*, Renaissance Classics, New York, 2012.

[58] D. Main, How hurricane forecasts have improved, *LiveScience* (2014), available at: http://www.livescience.com/21850-hurricane-forecast-improvements.html.

[59] S. Makridakis, A survey of time series, *International Statistical Review* **44**(1) (1976), 20–70.

[60] S. Makridakis, *Forecasting, Planning and Strategy for the 21st Century*, Free Press, New York, 1990.

[61] S. Makridakis et al., The accuracy of extrapolation (time series) methods: Results of a forecasting competition, *Journal of Forecasting* **1**(2) (1982), 111–153.

[62] S. Makridakis and M. Hibon, Accuracy of forecasting: An empirical investigation, *Journal of the Royal Statistical Society, Series A* **142**(2) (1979), 79–145 (with discussion).

[63] S. Makridakis and M. Hibon, The M3-competition: Results, conclusions and implications, *International Journal of Forecasting* **16** (2000), 451–476.

[64] S. Makridakis, R. Hogarth and A. Gaba, Forecasting and uncertainty in the economic and business world, *International Journal of Forecasting* **25**(4) (2009), 794–812.

[65] S. Makridakis, R. Hogarth and A. Gaba, *Dance with Chance: Making Luck Work for You*, Oneworld Publications, London, 2009.

[66] S. Makridakis, R. Hogarth and A. Gaba, Why forecasts fail. What to do instead, *MIT: Sloan Management Review* **151**(2) (2009), 83–91.

[67] S. Makridakis and A. Moleskis, The costs and benefits of positive illusions, *Frontiers in Psychology* (2015). doi:10.3389/fpsyg.2015.00859.

[68] S. Makridakis and R. Winkler, Average of forecasts: Some empirical results, *Management Science* **29**(9) (1983), 987–996.

[69] S. Makridakis and R. Winkler, Sampling distributions of post-sample forecasting errors, *Journal of the Royal Statistical Society* **38**(2) (1989), 331–342.

[70] B.G. Malkie, *A Random Walk Down Wall Street: The Time-Tested Strategy for Successful Investing*, 9th edn, Norton and Company, New York, 2007.

[71] G. McGregor, Climatology in support of climate risk management: A progress report, *Progress in Physical Geography* (2015). doi:10.1177/0309133315578941.

[72] F.P. McKenna, It won't happen to me: Unrealistic optimism or illusion of control?, *British Journal of Psychology* **84** (1993), 39–50. doi:10.1111/j.2044-8295.1993.tb02461.x.

[73] P. Meehl, *Clinical Versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence*, Univ. of Minnesota Press, Minneapolis, MN, 1954.

[74] A. Mehrotra, A.M. Zaslavsky and J.Z. Ayanian, Preventive health examinations and preventive gynecological examina-

tions in the United States, *Archives of Internal Medicine* **167**(17) (2007), 1876–1883.

[75] B. Mellers et al., Psychological strategies for winning a geopolitical forecasting tournament, *Psychological Science* **25**(5) (2014), 1106–1115.

[76] A.N.D. Meyer et al., Physicians' diagnostic accuracy, confidence, and resource requests, *JAMA Internal Medicine* **173**(21) (2013), 1952–1958.

[77] A.H. Murphy and R.L. Winkler, Probability forecasting in meteorology, *Journal of the American Statistical Association* **79**(387) (1984), 489–500.

[78] National Research Council, Committee on Abrupt Climate Change, Abrupt climate change: Inevitable surprises, 2002, available at: http://www.nap.edu/catalog/10136.html.

[79] T. Randall, These titans of oil are experts at making bold predictions: They just don't often come true, 2015, available at: http://www.bloomberg.com/news/articles/2015-04-29/these-titans-of-oil-are-experts-at-making-bold-predictions.

[80] C.F. Richter, A personal tribute, *Bulletin of the Seismological Society of America* **77**(6) (1987), 2234–2237.

[81] N. Silver, *The Signal and the Noise; Why so Many Predictions Fail – But Some Don't*, Penguin Books, New York, 2012.

[82] J. Simon, *The State of Humanity*, Blackwell, Cambridge, MA, 1997.

[83] J. Simon, *The Ultimate Resource 2*, Princeton Univ. Press, 1998.

[84] R.L. Smith, Extreme value theory based on the r largest annual events, *Journal of Hydrology* **86** (1986), 27–43.

[85] M. Sorenon, Prediction markets may revolutionize business forecasting, 2013, available at: http://globalriskinsights.com/2013/08/prediction-markets-may-revolutionise-business-forecasting/.

[86] D. Sornette and G. Ouillon, Dragon-kings: Mechanisms, statistical methods and empirical evidence, *Eur. Phys. J. Special Topics* **205** (2012), 1–26.

[87] J. Surowiecki, *The Wisdom of Crowds*, Anchor Books, New York, 2005.

[88] A.A. Syntetos, M.Z. Babai and E.S. Gardner Jr., Forecasting intermittent inventory demands: Simple parametric methods vs. bootstrapping, *Journal of Business* **68**(8) (2015), 1746–1752.

[89] N.N. Taleb, Errors, robustness, and the fourth quadrant, *International Journal of Forecasting* **25**(4) (2009), 744–759.

[90] N.N. Taleb, *Antifragile: Things That Gain from Disorder*, Random House, New York, 2012.

[91] N.N. Taleb, E. Canetti, T. Kinda, E. Loukoianova and C. Schmieder, A new heuristic measure of fragility and tail risks: Application to stress testing, IMF Working Paper WP/12/216, 2012.

[92] N.N. Taleb, R. Read, R. Douady, J. Norman and Y. Bar-Yam, The precautionary principle (with application to the genetic modification of organisms), *Extreme Risk Initiative – NYU School of Engineering Working Paper Series* (4 September 2014).

[93] P.E. Tetlock et al., GoodJudgement.com, 2015, available at: http://www.goodjudgment.com.

[94] The New York Times, How many mutual funds routinely rout the market? Zero, 2015.

[95] Theophrastus, *Theophrastus of Eresus on Winds and on Weather Signs*, G.J. Symons, ed., Edward Stanford, London, 1894. (Translated by J.G. Wood.)

[96] USGS, Top Earthquake States, 2012, available at: http://earthquake.usgs.gov/earthquakes/states/top_states.php.

[97] USGS, Earthquakes Hazard Program, 2015, available at: http://earthquake.usgs.gov/learn/topics/animal_eqs.php.

[98] USGS, Earthquakes Hazard Program, 2015, available at: http://earthquake.usgs.gov/learn/topics/groundwater.php.

[99] E. Zolfagharifard, Global warming believers are like a hysterical 'cult': MIT scientist compares 'climate alarmists' to religious fanatics, 2015, available at: http://www.dailymail.co.uk/sciencetech/article-2922553/Global-warming-believers-like-hysterical-cult-MIT-scientist-compares-climate-alarmists-religious-fanatics.html.