# The Signal and the Noise

*Why So Many Predictions Fail—but Some Don't*

NATE SILVER

THE PENGUIN PRESS
NEW YORK
2012

THE PENGUIN PRESS
Published by the Penguin Group
Penguin Group (USA) Inc., 375 Hudson Street,
New York, New York 10014, U.S.A.
Penguin Group (Canada), 90 Eglinton Avenue East, Suite 700, Toronto, Ontario, Canada M4P 2Y3 (a division of Pearson Penguin Canada Inc.) • Penguin Books Ltd, 80 Strand, London WC2R 0RL, England • Penguin Ireland, 25 St. Stephen's Green, Dublin 2, Ireland (a division of Penguin Books Ltd) • Penguin Books Australia Ltd, 250 Camberwell Road, Camberwell, Victoria 3124, Australia (a division of Pearson Australia Group Pty Ltd) • Penguin Books India Pvt Ltd, 11 Community Centre, Panchsheel Park, New Delhi – 110 017, India • Penguin Group (NZ), 67 Apollo Drive, Rosedale, Auckland 0632, New Zealand (a division of Pearson New Zealand Ltd) • Penguin Books (South Africa) (Pty) Ltd, 24 Sturdee Avenue, Rosebank, Johannesburg 2196, South Africa

Penguin Books Ltd, Registered Offices:
80 Strand, London WC2R 0RL, England

First published in 2012 by The Penguin Press,
a member of Penguin Group (USA) Inc.



Illustration credits
Figure 4-2: Courtesy of Dr. Tim Parker, University of Oxford
Figure 7-1: From "1918 Influenza: The Mother of All Pandemics" by Jeffery Taubenberger and David Morens, *Emerging Infectious Disease Journal*, vol. 12, no. 1, January 2006, Centers for Disease Control and Prevention
Figures 9-2, 9-3A, 9-3C, 9-4, 9-5, 9-6 and 9-7: By Cburnett, Wikimedia Commons
Figure 12-2: Courtesy of Dr. J. Scott Armstrong, The Wharton School, University of Pennsylvania

LIBRARY OF CONGRESS CATALOGING IN PUBLICATION DATA

Silver, Nate.
The signal and the noise : why most predictions fail but some don't / Nate Silver.
p. cm.
Includes bibliographical references and index.
ISBN 978-1-59420-411-1
1. Forecasting. 2. Forecasting—Methodology. 3. Forecasting—History.
4. Bayesian statistical decision theory. 5. Knowledge, Theory of. I. Title.
CB158.S54 2012
519.5'42—dc23 2012027308

Printed in the United States of America
3 5 7 9 10 8 6 4

DESIGNED BY AMANDA DEWEY



*To Mom and Dad*

# CONTENTS

*Introduction* 1

1. A CATASTROPHIC FAILURE OF PREDICTION 19
2. ARE YOU SMARTER THAN A TELEVISION PUNDIT? 47
3. ALL I CARE ABOUT IS W'S AND L'S 74
4. FOR YEARS YOU'VE BEEN TELLING US THAT RAIN IS GREEN 108
5. DESPERATELY SEEKING SIGNAL 142
6. HOW TO DROWN IN THREE FEET OF WATER 176
7. ROLE MODELS 204
8. LESS AND LESS AND LESS WRONG 232
9. RAGE AGAINST THE MACHINES 262
10. THE POKER BUBBLE 294
11. IF YOU CAN'T BEAT 'EM . . . 329
12. A CLIMATE OF HEALTHY SKEPTICISM 370
13. WHAT YOU DON'T KNOW CAN HURT YOU 412

*Conclusion* 446
*Acknowledgments* 455
*Notes* 459
*Index* 515

# INTRODUCTION

This is a book about information, technology, and scientific progress. This is a book about competition, free markets, and the evolution of ideas. This is a book about the things that make us smarter than any computer, and a book about human error. This is a book about how we learn, one step at a time, to come to knowledge of the objective world, and why we sometimes take a step back.

This is a book about prediction, which sits at the intersection of all these things. It is a study of why some predictions succeed and why some fail. My hope is that we might gain a little more insight into planning our futures and become a little less likely to repeat our mistakes.

## More Information, More Problems

The original revolution in information technology came not with the microchip, but with the printing press. Johannes Gutenberg's invention in 1440 made information available to the masses, and the explosion of ideas it produced had unintended consequences and unpredictable effects. It was a spark for the Industrial Revolution in 1775,[1] a tipping point in which civilization suddenly went from having made almost no scientific or economic progress for most of its existence to the exponential rates of growth and change that are familiar to us today. It set in motion the events that would produce the European Enlightenment and the founding of the American Republic.

But the printing press would first produce something else: hundreds of years of holy war. As mankind came to believe it could predict its fate and choose its destiny, the bloodiest epoch in human history followed.[2]

Books had existed prior to Gutenberg, but they were not widely written and they were not widely read. Instead, they were luxury items for the nobility, produced one copy at a time by scribes.[3] The going rate for reproducing a single manuscript was about one florin (a gold coin worth about $200 in today's dollars) per five pages,[4] so a book like the one you're reading now would cost around $20,000. It would probably also come with a litany of transcription errors, since it would be a copy of a copy of a copy, the mistakes having multiplied and mutated through each generation.

This made the accumulation of knowledge extremely difficult. It required heroic effort to prevent the volume of recorded knowledge from actually *decreasing*, since the books might decay faster than they could be reproduced. Various editions of the Bible survived, along with a small number of canonical texts, like from Plato and Aristotle. But an untold amount of wisdom was lost to the ages,[5] and there was little incentive to record more of it to the page.

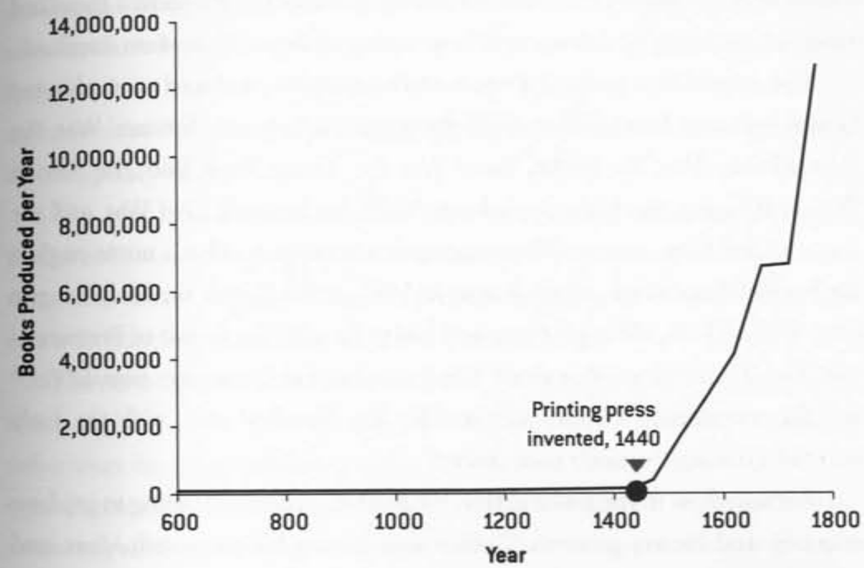
The pursuit of knowledge seemed inherently futile, if not altogether vain. If today we feel a sense of impermanence because things are changing so rapidly, impermanence was a far more literal concern for the generations before us. There was "nothing new under the sun," as the beautiful Bible verses in Ecclesiastes put it—not so much because everything had been discovered but because everything would be forgotten.[6]

The printing press changed that, and did so permanently and profoundly. Almost overnight, the cost of producing a book decreased by about three hundred times,[7] so a book that might have cost $20,000 in today's dollars instead cost $70. Printing presses spread very rapidly throughout Europe; from Gutenberg's Germany to Rome, Seville, Paris, and Basel by 1470, and then to almost all other major European cities within another ten years.[8] The number of books being produced grew exponentially, increasing by about thirty times in the first century after the printing press was invented.[9] The store of human knowledge had begun to accumulate, and rapidly.





FIGURE I-1: EUROPEAN BOOK PRODUCTION




As was the case during the early days of the World Wide Web, however, the quality of the information was highly varied. While the printing press paid almost immediate dividends in the production of higher quality maps,[10] the bestseller list soon came to be dominated by heretical religious texts and pseudoscientific ones.[11] Errors could now be mass-produced, like in the so-called Wicked Bible, which committed the most unfortunate typo in history to the page: thou *shalt* commit adultery.[12] Meanwhile, exposure to so many new ideas was producing mass confusion. The amount of information was increasing much more rapidly than our understanding of what to do with it, or our ability to differentiate the useful information from the mistruths.[13] Paradoxically, the result of having so much more shared knowledge was increasing isolation along national and religious lines. The instinctual shortcut that we take when we have "too much information" is to engage with it selectively, picking out the parts we like and ignoring the remainder, making allies with those who have made the same choices and enemies of the rest.

The most enthusiastic early customers of the printing press were those who used it to evangelize. Martin Luther's *Ninety-five Theses* were not that radical; similar sentiments had been debated many times over. What was revolutionary, as Elizabeth Eisenstein writes, is that Luther's theses "did not stay tacked to the church door."[14] Instead, they were reproduced at least three hundred thousand times by Gutenberg's printing press[15]—a runaway hit even by modern standards.

The schism that Luther's Protestant Reformation produced soon plunged Europe into war. From 1524 to 1648, there was the German Peasants' War, the Schmalkaldic War, the Eighty Years' War, the Thirty Years' War, the French Wars of Religion, the Irish Confederate Wars, the Scottish Civil War, and the English Civil War—many of them raging simultaneously. This is not to neglect the Spanish Inquisition, which began in 1480, or the War of the Holy League from 1508 to 1516, although those had less to do with the spread of Protestantism. The Thirty Years' War alone killed one-third of Germany's population,[16] and the seventeenth century was possibly the bloodiest ever, with the early twentieth staking the main rival claim.[17]

But somehow in the midst of this, the printing press was starting to produce scientific and literary progress. Galileo was sharing his (censored) ideas, and Shakespeare was producing his plays.

Shakespeare's plays often turn on the idea of fate, as much drama does. What makes them so tragic is the gap between what his characters might like to accomplish and what fate provides to them. The idea of controlling one's fate seemed to have become part of the human consciousness by Shakespeare's time—but not yet the competencies to achieve that end. Instead, those who tested fate usually wound up dead.[18]

These themes are explored most vividly in *The Tragedy of Julius Caesar.* Throughout the first half of the play Caesar receives all sorts of apparent warning signs—what he calls predictions[19] ("beware the ides of March")—that his coronation could turn into a slaughter. Caesar of course ignores these signs, quite proudly insisting that they point to someone else's death—or otherwise reading the evidence selectively. Then Caesar is assassinated.





"[But] men may construe things after their fashion / Clean from the purpose of the things themselves," Shakespeare warns us through the voice of Cicero—good advice for anyone seeking to pluck through their newfound wealth of information. It was hard to tell the signal from the noise. The story the data tells us is often the one we'd like to hear, and we usually make sure that it has a happy ending.

And yet if *The Tragedy of Julius Caesar* turned on an ancient idea of prediction—associating it with fatalism, fortune-telling, and superstition—it also introduced a more modern and altogether more radical idea: that we might interpret these signs so as to gain an advantage from them. "Men at some time are masters of their fates," says Cassius, hoping to persuade Brutus to partake in the conspiracy against Caesar.

The idea of man as master of his fate was gaining currency. The words *predict* and *forecast* are largely used interchangeably today, but in Shakespeare's time, they meant different things. A prediction was what the soothsayer told you; a forecast was something more like Cassius's idea.

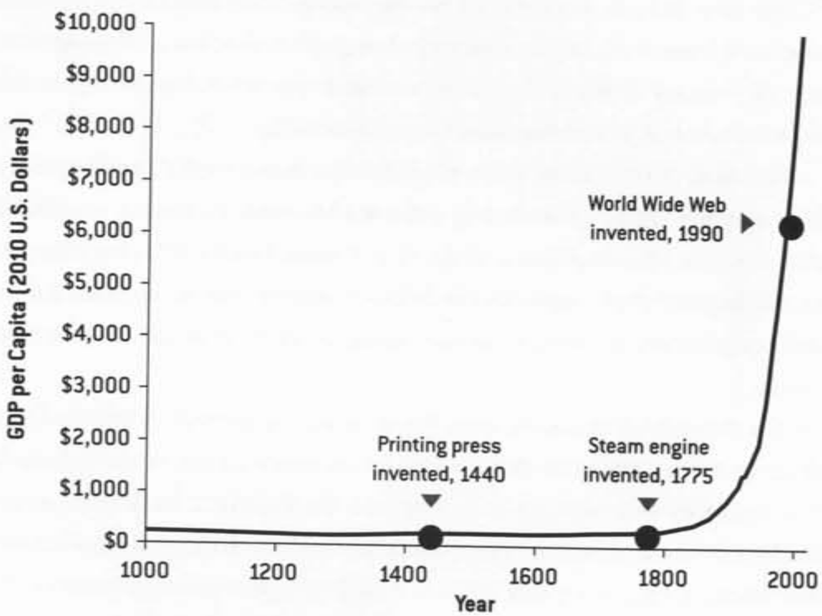
The term *forecast* came from English's Germanic roots,[20] unlike *predict*, which is from Latin.[21] Forecasting reflected the new Protestant worldliness rather than the otherworldliness of the Holy Roman Empire. Making a forecast typically implied planning under conditions of uncertainty. It suggested having prudence, wisdom, and industriousness, more like the way we now use the word *foresight*.[22]

The theological implications of this idea are complicated.[23] But they were less so for those hoping to make a gainful existence in the terrestrial world. These qualities were strongly associated with the Protestant work ethic, which Max Weber saw as bringing about capitalism and the Industrial Revolution.[24] This notion of *forecasting* was very much tied in to the notion of progress. All that information in all those books ought to have helped us to plan our lives and profitably predict the world's course.

The Protestants who ushered in centuries of holy war were learning how to use their accumulated knowledge to change society. The Industrial Revolution largely began in Protestant countries and largely in those with a free press, where both religious and scientific ideas could flow without fear of censorship.[25]

The importance of the Industrial Revolution is hard to overstate. Throughout essentially all of human history, economic growth had proceeded at a rate of perhaps 0.1 percent per year, enough to allow for a very gradual increase in population, but not *any* growth in per capita living standards.[26] And then, suddenly, there was progress when there had been none. Economic growth began to zoom upward much faster than the growth rate of the population, as it has continued to do through to the present day, the occasional global financial meltdown notwithstanding.[27]

FIGURE I-2: GLOBAL PER CAPITA GDP, 1000–2010




The explosion of information produced by the printing press had done us a world of good, it turned out. It had just taken 330 years—and millions dead in battlefields around Europe—for those advantages to take hold.





## The Productivity Paradox

We face danger whenever information growth outpaces our understanding of how to process it. The last forty years of human history imply that it can still take a long time to translate information into useful knowledge, and that if we are not careful, we may take a step back in the meantime.

The term "information age" is not particularly new. It started to come into more widespread use in the late 1970s. The related term "computer age" was used earlier still, starting in about 1970.[28] It was at around this time that computers began to be used more commonly in laboratories and academic settings, even if they had not yet become common as home appliances. This time it did not take three hundred years before the growth in information technology began to produce tangible benefits to human society. But it did take fifteen to twenty.

The 1970s were the high point for "vast amounts of theory applied to extremely small amounts of data," as Paul Krugman put it to me. We had begun to use computers to produce models of the world, but it took us some time to recognize how crude and assumption laden they were, and that the precision that computers were capable of was no substitute for predictive accuracy. In fields ranging from economics to epidemiology, this was an era in which bold predictions were made, and equally often failed. In 1971, for instance, it was claimed that we would be able to predict earthquakes within a decade,[29] a problem that we are no closer to solving forty years later.

Instead, the computer boom of the 1970s and 1980s produced a temporary *decline* in economic and scientific productivity. Economists termed this the productivity paradox. "You can see the computer age everywhere but in the productivity statistics," wrote the economist Robert Solow in 1987.[30] The United States experienced four distinct recessions between 1969 and 1982.[31] The late 1980s were a stronger period for our economy, but less so for countries elsewhere in the world.

Scientific progress is harder to measure than economic progress.[32] But one mark of it is the number of patents produced, especially relative to the investment in research and development. If it has become cheaper to produce a new invention, this suggests that we are using our information wisely and are forging it into knowledge. If it is becoming more expensive, this suggests that we are seeing signals in the noise and wasting our time on false leads.

In the 1960s the United States spent about $1.5 million (adjusted for inflation[33]) per patent application[34] by an American inventor. That figure *rose* rather than fell at the dawn of the information age, however, doubling to a peak of about $3 million in 1986.[35]

FIGURE I-3: RESEARCH AND DEVELOPMENT EXPENDITURES PER PATENT APPLICATION


As we came to more realistic views of what that new technology could accomplish for us, our research productivity began to improve again in the 1990s. We wandered up fewer blind alleys; computers began to improve our everyday lives and help our economy. Stories of prediction are often those of long-term progress but short-term regress. Many things that seem predictable over the long run foil our best-laid plans in the meanwhile.





## The Promise and Pitfalls of "Big Data"

The fashionable term now is "Big Data." IBM estimates that we are generating 2.5 quintillion bytes of data each day, more than 90 percent of which was created in the last two years.[36]

This exponential growth in information is sometimes seen as a cure-all, as computers were in the 1970s. Chris Anderson, the editor of *Wired* magazine, wrote in 2008 that the sheer volume of data would obviate the need for theory, and even the scientific method.[37]

This is an emphatically pro-science and pro-technology book, and I think of it as a very optimistic one. But it argues that these views are badly mistaken. The numbers have no way of speaking for themselves. We speak for them. We imbue them with meaning. Like Caesar, we may construe them in self-serving ways that are detached from their objective reality.

Data-driven predictions can succeed—and they can fail. It is when we deny our role in the process that the odds of failure rise. Before we demand more of our data, we need to demand more of ourselves.

This attitude might seem surprising if you know my background. I have a reputation for working with data and statistics and using them to make successful predictions. In 2003, bored at a consulting job, I designed a system called PECOTA, which sought to predict the statistics of Major League Baseball players. It contained a number of innovations—its forecasts were probabilistic, for instance, outlining a range of possible outcomes for each player—and we found that it outperformed competing systems when we compared their results. In 2008, I founded the Web site FiveThirtyEight, which sought to forecast the upcoming election. The FiveThirtyEight forecasts correctly predicted the winner of the presidential contest in forty-nine of fifty states as well as the winner of all thirty-five U.S. Senate races.

After the election, I was approached by a number of publishers who wanted to capitalize on the success of books such as *Moneyball* and *Freakonomics* that told the story of nerds conquering the world. This book was conceived of along those lines—as an investigation of data-driven predictions in fields ranging from baseball to finance to national security.

But in speaking with well more than one hundred experts in more than a dozen fields over the course of four years, reading hundreds of journal articles and books, and traveling everywhere from Las Vegas to Copenhagen in pursuit of my investigation, I came to realize that prediction in the era of Big Data was not going very well. I had been lucky on a few levels: first, in having achieved success despite having made many of the mistakes that I will describe, and second, in having chosen my battles well.

Baseball, for instance, is an exceptional case. It happens to be an especially rich and revealing exception, and the book considers why this is so—why a decade after *Moneyball*, stat geeks and scouts are now working in harmony.

The book offers some other hopeful examples. Weather forecasting, which also involves a melding of human judgment and computer power, is one of them. Meteorologists have a bad reputation, but they have made remarkable progress, being able to forecast the landfall position of a hurricane three times more accurately than they were a quarter century ago. Meanwhile, I met poker players and sports bettors who really were beating Las Vegas, and the computer programmers who built IBM's Deep Blue and took down a world chess champion.

But these cases of progress in forecasting must be weighed against a series of failures.

If there is one thing that defines Americans—one thing that makes us exceptional—it is our belief in Cassius's idea that we are in control of our own fates. Our country was founded at the dawn of the Industrial Revolution by religious rebels who had seen that the free flow of ideas had helped to spread not just their religious beliefs, but also those of science and commerce. Most of our strengths and weaknesses as a nation—our ingenuity and our industriousness, our arrogance and our impatience—stem from our unshakable belief in the idea that we choose our own course.

But the new millennium got off to a terrible start for Americans. We had not seen the September 11 attacks coming. The problem was not want of information. As had been the case in the Pearl Harbor attacks six decades earlier, all the signals were there. But we had not put them together. Lacking a proper





theory for how terrorists might behave, we were blind to the data and the attacks were an "unknown unknown" to us.

There also were the widespread failures of prediction that accompanied the recent global financial crisis. Our naïve trust in models, and our failure to realize how fragile they were to our choice of assumptions, yielded disastrous results. On a more routine basis, meanwhile, I discovered that we are unable to predict recessions more than a few months in advance, and not for lack of trying. While there has been considerable progress made in controlling inflation, our economic policy makers are otherwise flying blind.

The forecasting models published by political scientists in advance of the 2000 presidential election predicted a landslide 11-point victory for Al Gore.[38] George W. Bush won instead. Rather than being an anomalous result, failures like these have been fairly common in political prediction. A long-term study by Philip E. Tetlock of the University of Pennsylvania found that when political scientists claimed that a political outcome had absolutely *no* chance of occurring, it nevertheless happened about 15 percent of the time. (The political scientists are probably better than television pundits, however.)

There has recently been, as in the 1970s, a revival of attempts to predict earthquakes, most of them using highly mathematical and data-driven techniques. But these predictions envisaged earthquakes that never happened and failed to prepare us for those that did. The Fukushima nuclear reactor had been designed to handle a magnitude 8.6 earthquake, in part because some seismologists concluded that anything larger was impossible. Then came Japan's horrible magnitude 9.1 earthquake in March 2011.

There are entire disciplines in which predictions have been failing, often at great cost to society. Consider something like biomedical research. In 2005, an Athens-raised medical researcher named John P. Ioannidis published a controversial paper titled "Why Most Published Research Findings Are False."[39] The paper studied positive findings documented in peer-reviewed journals: descriptions of successful predictions of medical hypotheses carried out in laboratory experiments. It concluded that most of these findings were likely to fail when applied in the real world. Bayer Laboratories recently confirmed Ioannidis's hypothesis. They could not replicate about *two-thirds* of the positive findings claimed in medical journals when they attempted the experiments themselves.[40]

Big Data *will* produce progress—eventually. How quickly it does, and whether we regress in the meantime, will depend on us.

## Why the Future Shocks Us

Biologically, we are not very different from our ancestors. But some stone-age strengths have become information-age weaknesses.

Human beings do not have very many natural defenses. We are not all that fast, and we are not all that strong. We do not have claws or fangs or body armor. We cannot spit venom. We cannot camouflage ourselves. And we cannot fly. Instead, we survive by means of our wits. Our minds are quick. We are wired to detect patterns and respond to opportunities and threats without much hesitation.

"This need of finding patterns, humans have this more than other animals," I was told by Tomaso Poggio, an MIT neuroscientist who studies how our brains process information. "Recognizing objects in difficult situations means generalizing. A newborn baby can recognize the basic pattern of a face. It has been learned by evolution, not by the individual."

The problem, Poggio says, is that these evolutionary instincts sometimes lead us to see patterns when there are none there. "People have been doing that all the time," Poggio said. "Finding patterns in random noise."

The human brain is quite remarkable; it can store perhaps three terabytes of information.[41] And yet that is only about one one-millionth of the information that IBM says is now produced in the world *each day*. So we have to be terribly selective about the information we choose to remember.

Alvin Toffler, writing in the book *Future Shock* in 1970, predicted some of the consequences of what he called "information overload." He thought our defense mechanism would be to simplify the world in ways that confirmed our biases, even as the world itself was growing more diverse and more complex.[42]

Our biological instincts are not always very well adapted to the information-rich modern world. Unless we work *actively* to become aware of the biases





we introduce, the returns to additional information may be minimal—or diminishing.

The information overload after the birth of the printing press produced greater sectarianism. Now those different religious ideas could be testified to with more information, more conviction, more "proof"—and less tolerance for dissenting opinion. The same phenomenon seems to be occurring today. Political partisanship began to increase very rapidly in the United States beginning at about the time that Tofller wrote *Future Shock* and it may be accelerating even faster with the advent of the Internet.[43]

These partisan beliefs can upset the equation in which more information will bring us closer to the truth. A recent study in *Nature* found that the *more* informed that strong political partisans were about global warming, the *less* they agreed with one another.[44]

Meanwhile, if the quantity of information is increasing by 2.5 quintillion bytes per day, the amount of *useful* information almost certainly isn't. Most of it is just noise, and the noise is increasing faster than the signal. There are so many hypotheses to test, so many data sets to mine—but a relatively constant amount of objective truth.

The printing press changed the way in which we made mistakes. Routine errors of transcription became less common. But when there was a mistake, it would be reproduced many times over, as in the case of the Wicked Bible.

Complex systems like the World Wide Web have this property. They may not fail as often as simpler ones, but when they fail they fail badly. Capitalism and the Internet, both of which are incredibly efficient at propagating information, create the potential for bad ideas as well as good ones to spread. The bad ideas may produce disproportionate effects. In advance of the financial crisis, the system was so highly levered that a single lax assumption in the credit ratings agencies' models played a huge role in bringing down the whole global financial system.

Regulation is one approach to solving these problems. But I am suspicious that it is an excuse to avoid looking within ourselves for answers. We need to stop, and admit it: we have a prediction problem. We love to predict things—and we aren't very good at it.

## The Prediction Solution

If prediction is the central problem of this book, it is also its solution.

Prediction is indispensable to our lives. Every time we choose a route to work, decide whether to go on a second date, or set money aside for a rainy day, we are making a forecast about how the future will proceed—and how our plans will affect the odds for a favorable outcome.

Not all of these day-to-day problems require strenuous thought; we can budget only so much time to each decision. Nevertheless, you are making predictions many times every day, whether or not you realize it.

For this reason, this book views prediction as a shared enterprise rather than as a function that a select group of experts or practitioners perform. It is amusing to poke fun at the experts when their predictions fail. However, we should be careful with our Schadenfreude. To say our predictions are no worse than the experts' is to damn ourselves with some awfully faint praise.

Prediction does play a particularly important role in science, however. Some of you may be uncomfortable with a premise that I have been hinting at and will now state explicitly: we can *never* make perfectly objective predictions. They will *always* be tainted by our subjective point of view.

But this book is emphatically against the nihilistic viewpoint that there is no objective truth. It asserts, rather, that a belief in the objective truth—and a commitment to pursuing it—is the first prerequisite of making better predictions. The forecaster's next commitment is to realize that she perceives it imperfectly.

Prediction is important because it connects subjective and objective reality. Karl Popper, the philosopher of science, recognized this view.[45] For Popper, a hypothesis was not scientific unless it was falsifiable—meaning that it could be tested in the real world by means of a prediction.

What should give us pause is that the few ideas we have tested aren't doing so well, and many of our ideas have not or cannot be tested at all. In economics, it is much easier to test an unemployment rate forecast than a claim about the effectiveness of stimulus spending. In political science, we can test





models that are used to predict the outcome of elections, but a theory about how changes to political institutions might affect policy outcomes could take decades to verify.

I do not go as far as Popper in asserting that such theories are therefore unscientific or that they lack any value. However, the fact that the few theories we *can* test have produced quite poor results suggests that many of the ideas we *haven't* tested are very wrong as well. We are undoubtedly living with many delusions that we do not even realize.

But there is a way forward. It is not a solution that relies on half-baked policy ideas—particularly given that I have come to view our political system as a big part of the problem. Rather, the solution requires an attitudinal change.

This attitude is embodied by something called Bayes's theorem, which I introduce in chapter 8. Bayes's theorem is nominally a mathematical formula. But it is really much more than that. It implies that we must think differently about our ideas—and how to test them. We must become more comfortable with probability and uncertainty. We must think more carefully about the assumptions and beliefs that we bring to a problem.

The book divides roughly into halves. The first seven chapters diagnose the prediction problem while the final six explore and apply Bayes's solution.

Each chapter is oriented around a particular subject and describes it in some depth. There is no denying that this is a detailed book—in part because that is often where the devil lies, and in part because my view is that a certain amount of immersion in a topic will provide disproportionately more insight than an executive summary.

The subjects I have chosen are usually those in which there is some publicly shared information. There are fewer examples of forecasters making predictions based on private information (for instance, how a company uses its customer records to forecast demand for a new product). My preference is for topics where you can check out the results for yourself rather than having to take my word for it.

## A Short Road Map to the Book

The book weaves between examples from the natural sciences, the social sciences, and from sports and games. It builds from relatively straightforward cases, where the successes and failures of prediction are more easily demarcated, into others that require slightly more finesse.

Chapters 1 through 3 consider the failures of prediction surrounding the recent financial crisis, the successes in baseball, and the realm of political prediction—where some approaches have worked well and others haven't. They should get you thinking about some of the most fundamental questions that underlie the prediction problem. How can we apply our judgment to the data—without succumbing to our biases? When does market competition make forecasts better—and how can it make them worse? How do we reconcile the need to use the past as a guide with our recognition that the future may be different?

Chapters 4 through 7 focus on *dynamic* systems: the behavior of the earth's atmosphere, which brings about the weather; the movement of its tectonic plates, which can cause earthquakes; the complex human interactions that account for the behavior of the American economy; and the spread of infectious diseases. These systems are being studied by some of our best scientists. But dynamic systems make forecasting more difficult, and predictions in these fields have not always gone very well.

Chapters 8 through 10 turn toward solutions—first by introducing you to a sports bettor who applies Bayes's theorem more expertly than many economists or scientists do, and then by considering two other games, chess and poker. Sports and games, because they follow well-defined rules, represent good laboratories for testing our predictive skills. They help us to a better understanding of randomness and uncertainty and provide insight about how we might forge information into knowledge.

Bayes's theorem, however, can also be applied to more existential types of problems. Chapters 11 through 13 consider three of these cases: global warming, terrorism, and bubbles in financial markets. These are hard problems for forecasters and for society. But if we are up to the challenge, we can make our country, our economy, and our planet a little safer.





. . .

The world has come a long way since the days of the printing press. Information is no longer a scarce commodity; we have more of it than we know what to do with. But relatively little of it is useful. We perceive it selectively, subjectively, and without much self-regard for the distortions that this causes. We think we want information when we really want knowledge.

The signal is the truth. The noise is what distracts us from the truth. This is a book about the signal and the noise.