FEBRUARY 2014
McKinsey Quarterly

# The benefits—and limits—of decision models

**Phil Rosenzweig**

Big data and models help overcome biases that cloud judgment, but many executive decisions also require bold action inspired by self-confidence. Here's how to take charge in a clear-headed way.

**The growing power** of decision models has captured plenty of C-suite attention in recent years. Combining vast amounts of data and increasingly sophisticated algorithms, modeling has opened up new pathways for improving corporate performance.[1] Models can be immensely useful, often making very accurate predictions or guiding knotty optimization choices and, in the process, can help companies to avoid some of the common biases that at times undermine leaders' judgments.

Yet when organizations embrace decision models, they sometimes overlook the need to use them well. In this article, I'll address an important distinction between outcomes leaders can influence and those they cannot. For things that executives cannot directly influence, accurate judgments are paramount and the new modeling tools can be valuable. However, when a senior manager can have a direct influence over the outcome of a decision, the challenge is quite different. In this case, the task isn't to *predict* what will happen but

[1] See these recent articles: Brad Brown, Michael Chui, and James Manyika, "Are you ready for the era of 'big data'?," *McKinsey Quarterly*, October 2011; Erik Brynjolfsson, Jeff Hammerbacher, and Brad Stevens, "Competing through data: Three experts offer their game plans," *McKinsey Quarterly*, October 2011; Brad Brown, David Court, and Paul Willmott, "Mobilizing your C-suite for big-data analytics," *McKinsey Quarterly*, November 2013; and Stefan Biesdorf, David Court, and Paul Willmott, "Big data: What's your plan?," *McKinsey Quarterly*, March 2013, all available on mckinsey.com.

to *make* it happen. Here, positive thinking—indeed, a healthy dose of management confidence—can make the difference between success and failure.

## Where models work well

Examples of successful decision models are numerous and growing. Retailers gather real-time information about customer behavior by monitoring preferences and spending patterns. They can also run experiments to test the impact of changes in pricing or packaging and then rapidly observe the quantities sold. Banks approve loans and insurance companies extend coverage, basing their decisions on models that are continually updated, factoring in the most information to make the best decisions.

Some recent applications are truly dazzling. Certain companies analyze masses of financial transactions in real time to detect fraudulent credit-card use. A number of companies are gathering years of data about temperature and rainfall across the United States to run weather simulations and help farmers decide what to plant and when. Better risk management and improved crop yields are the result.

Other examples of decision models border on the humorous. Garth Sundem and John Tierney devised a model to shed light on what they described, tongues firmly in cheek, as one of the world's great unsolved mysteries: how long will a celebrity marriage last? They came up with the Sundem/Tierney Unified Celebrity Theory, which predicted the length of a marriage based on the couple's combined age (older was better), whether either had tied the knot before (failed marriages were not a good sign), and how long they had dated (the longer the better). The model also took into account fame (measured by hits on a Google search) and sex appeal (the share of those Google hits that came up with images of the wife scantily clad). With only a handful of variables, the model did a very good job of predicting the fate of celebrity marriages over the next few years.

Models have also shown remarkable power in fields that are usually considered the domain of experts. With data from France's premier wine-producing regions, Bordeaux and Burgundy, Princeton economist Orley Ashenfelter devised a model that used just three variables to predict the quality of a vintage: winter rainfall, harvest rainfall, and average growing-season temperature. To the surprise of many, the model outperformed wine connoisseurs.

Why do decision models perform so well? In part because they can gather vast quantities of data, but also because they avoid common biases that undermine human judgment.[2] People tend to be overly precise, believing that their estimates will be more accurate than they really are. They suffer from the recency bias, placing too much weight on the most immediate information. They are also unreliable: ask someone the same question on two different occasions and you may get two different answers. Decision models have none of these drawbacks; they weigh all data objectively and evenly. No wonder they do better than humans.

## Can we control outcomes?

With so many impressive examples, we might conclude that decision models can improve just about anything. That would be a mistake. Executives need not only to appreciate the power of models but also to be cognizant of their limits.

Look back over the previous examples. In every case, the goal was to make a prediction about something that could not be influenced directly. Models can estimate whether a loan will be repaid but won't actually change the likelihood that payments will arrive on time, give borrowers a greater capacity to pay, or make sure they don't squander their money before payment is due. Models can predict the rainfall and days of sunshine on a given farm in central Iowa but can't change the weather. They can estimate how long a celebrity marriage might last but won't help it last longer or cause another to end sooner. They can predict the quality of a wine vintage but won't make the wine any better, reduce its acidity, improve the balance, or change the undertones. For these sorts of estimates, finding ways to avoid bias and maintain accuracy is essential.

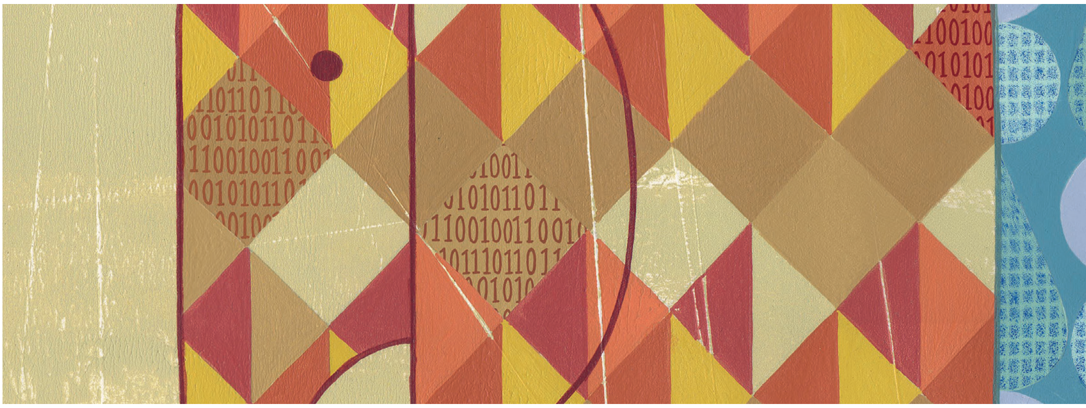
[2] Dan P. Lovallo and Olivier Sibony, "Distortions and deceptions in strategic decisions," *McKinsey Quarterly*, February 2006, mckinsey.com.

Executives, however, are not concerned only with predicting things they cannot influence. Their primary duty—as the word *execution* implies—is to get things done. The task of leadership is to mobilize people to achieve a desired end. For that, leaders need to inspire their followers to reach demanding goals, perhaps even to do more than they have done before or believe is possible. Here, positive thinking matters. Holding a somewhat exaggerated level of self-confidence isn't a dangerous bias; it often helps to stimulate higher performance.

This distinction seems simple but it's often overlooked. In our embrace of decision models, we sometimes forget that so much of life is about getting things done, not predicting things we cannot control.

## The insight of Billy Beane . . .

The failure to distinguish between what we can and cannot control has led to confusion in many fields. Perhaps nowhere has the gap been more evident than in the application of decision models to baseball. For decades, baseball managers made tactical decisions according to an unwritten set of rules, known as *going by the book*. Beginning in the 1970s, a group of statistically minded fans—practitioners of *sabermetrics*, a term coined in honor of the Society for American Baseball Research—began to apply the power of data analysis to test some of baseball's cherished notions, often with surprising results.



Take a common tactic, the sacrifice bunt. With a runner on first base and one or no outs, should the batter bunt the ball to advance the runner? Conventional wisdom said yes. As Bill James, a pioneer of sabermetrics, put it, "The experts all knew that when there was a runner on first and no one out, the percentage move was to bunt."[3] Until recently, there was no way to conduct a decent empirical analysis of the sacrifice bunt, but now there is. A simple test compares the runs that result from two situations: a runner at first base with no outs and a runner at second base with one out. Analyzing an entire season of major-league games revealed that, on average, making an out to advance the runner leads to fewer runs. The sacrifice bunt is just one example of how conventional wisdom in baseball can be wrong. James concluded, "A very, very large percentage of the things that the experts all knew to be true turned out, on examination, not to be true at all."[4]

The use of data analysis was the key insight of Michael Lewis's 2003 bestseller, *Moneyball: The Art of Winning an Unfair Game*. Lewis described how the Oakland Athletics, a low-budget team in a small market, posted several consecutive years of excellent results. Athletics general manager Billy Beane used decision models to discover what truly led to a winning performance and applied those insights to assemble a team of very good players at bargain prices. In "decision speak," he was trying to *optimize runs scored per dollar spent*. Oakland compiled a strong record for several consecutive years, despite a low payroll, largely thanks to its reliance on decision analytics.

With the publication of *Moneyball*, the use of statistics in baseball became widely accepted. Statistically minded general managers, some of them disciples of Billy Beane, spread throughout major-league baseball. Soon a host of new statistics was devised to measure increasingly esoteric aspects of play. One tracks the location and velocity of every single pitch, providing for the ever-finer analysis of any pitcher's performance. Another records every ball in play and extends statistical analysis to fielding, the aspect of play least amenable to quantification. Insights into the batting patterns of

[3] Bill James, *Solid Fool's Gold: Detours on the Way to Conventional Wisdom*, first edition, Chicago, IL: ACTA Sports, 2011, p. 185.

[4] James, p. 186.

individual players now help teams to shift the positions of their fielders for each batter. America's pastime has fully embraced the digital age, bringing Cupertino to Cooperstown.

## . . . and the wisdom of Joe Morgan

The notion that players could be evaluated by statistical models was not universally accepted. Players, in particular, insisted that performance couldn't be reduced to figures. Statistics don't capture the intangibles of the game, they argued, or grasp the subtle qualities that make players great. Of all the critics, none was more outspoken than Joe Morgan, a star player from the 1960s and 1970s. "I don't think that statistics are what the game is about," Morgan insisted. "I played the Game. I know what happens out there. . . . *Players* win games. Not theories."[5]

Proponents of statistical analysis dismissed Joe Morgan as unwilling to accept the truth, but in fact he wasn't entirely wrong. Models are useful in predicting things we cannot control, but for players—on the field and in the midst of a game—the reality is different. Players don't *predict* performance; they have to *achieve* it. For that purpose, impartial and dispassionate analysis is insufficient. Positive thinking matters, too.

When we stand back from the claims and counterclaims, Billy Beane and Joe Morgan are both right—just about different things. The job of a general manager is to assemble a team that will perform well on the field. When general managers evaluate players, decide whom to sign and how much to pay, whom to promote and whom to trade, they do best by relying on dispassionate analysis. There's nothing to be gained from wishful thinking or biased judgments. Billy Beane was known to work out in the clubhouse gym during games rather than watch the action on the diamond. Why? Because as general manager, he doesn't throw a ball or swing a bat. He can exercise control over the composition of the team, but once the game begins he's powerless.

[5]Tommy Craggs, "Say-It-Ain't-So Joe," *SF Weekly*, July 6, 2005, sfweekly.com.

## Part of the appeal of decision models lies in their ability to make predictions, to compare those predictions with what actually happens, and then to evolve so as to make more accurate predictions.

For players, the reality is entirely different: their job is to hit the ball and drive in the runs. A mind-set with high self-confidence—even a level of confidence that is, by some definitions, slightly excessive—is vital. Perhaps it shouldn't be surprising that this point was articulated so intuitively by Joe Morgan, a diminutive man who not only won the National League's Most Valuable Player award in 1975 and 1976 but is also considered one of the greatest second basemen of all time.

Pitting baseball traditionalists against proponents of statistical analysis makes for a spirited debate. But that's a false dichotomy, not conducive to a better understanding of the game. When the *Moneyball* controversy was at its height, St. Louis Cardinals manager Tony LaRussa wisely observed that no single approach was best: "The 'Moneyball' kind of stuff has its place, but so does the human. Really, the combination is the answer."[6]

The same distinction applies to managers of all kinds. The question, boiled down to its essence, is whether we are trying to predict something we cannot influence or something we can control, at least in part. Decision models are increasingly powerful for tasks requiring the impartial analysis of vast amounts of data. When we can and must shape outcomes, however, they do not suffice. An executive may be wise to rely on decision models when estimating consumer reactions to a promotion or meteorological conditions, but motivating a team to achieve high performance is a different matter. A combination of skills is the answer.

[6] David Leonhardt, "Science and art at odds on the field of dreams," the *New York Times*, August 29, 2005, nytimes.com.

## Influence, direct and indirect

The use of decision models raises a third possibility, in addition to direct influence and no influence: *indirect influence*. Even if we cannot directly shape an outcome, a model's prediction may be communicated in a way that alters behavior and indirectly shapes an outcome. Indirect influence takes two forms. If it increases the chance an event will occur, that's a self-fulfilling prediction. If it lowers the chance an event will occur, that's a self-negating prediction.

Consider a bank that uses a decision model to review loan applications. The model has no direct influence on a borrower's behavior; it can't control spending habits or make sure that anyone saves enough money each month to repay a loan. Suppose that instead of simply turning down the application, however, a banker meets an aspiring borrower and explains the reasons for concern. Such an intervention could cause the applicant to behave differently, perhaps by devising a monthly budget or by asking for direct payment via payroll deduction. Even though the model had no direct influence on the outcome, it could exert an indirect influence. The same goes for one of the most impressive examples of decision models in recent years: the electoral model. Models do not directly affect the outcome of an election—they do not cast votes. But if the projections of models are communicated broadly, they may embolden some supporters and discourage others, and thereby have an indirect influence.

The crucial lesson for executives is not simply to marvel at the power of decision analytics but also to understand the role these techniques play in achieving a desired outcome. If that outcome is an accurate prediction, models have unparalleled power. If we can shape it, then concerted effort—aided by positive thinking—can be vital. And in some instances, the output of a decision model can be communicated to influence a desired outcome indirectly. Models are powerful tools; keeping in mind the desired end is paramount.

## Improving models over time

Part of the appeal of decision models lies in their ability to make predictions, to compare those predictions with what actually happens, and then to evolve so as to make more accurate predictions.

In retailing, for example, companies can run experiments with different combinations of price and packaging, then rapidly obtain feedback and alter their marketing strategy. Netflix captures rapid feedback to learn what programs have the greatest appeal and then uses those insights to adjust its offerings. Models are not only useful at any particular moment but can also be updated over time to become more and more accurate.

Using feedback to improve models is a powerful technique but is more applicable in some settings than in others. Dynamic improvement depends on two features: first, the observation of results should not make any future occurrence either more or less likely and, second, the feedback cycle of observation and adjustment should happen rapidly. Both conditions hold in retailing, where customer behavior can be measured without directly altering it and results can be applied rapidly, with prices or other features changed almost in real time. They also hold in weather forecasting, since daily measurements can refine models and help to improve subsequent predictions. The steady improvement of models that predict weather—from an average error (in the maximum temperature) of 6 degrees Fahrenheit in the early 1970s to 5 degrees in the 1990s and just 4 by 2010—is testimony to the power of updated models.

For other events, however, these two conditions may not be present. As noted, executives not only estimate things they cannot affect but are also charged with bringing about outcomes. Some of the most consequential decisions of all—including the launch of a new product, entry into a new market, or the acquisition of a rival—are about mobilizing resources to get things done. Furthermore, the results are not immediately visible and may take months or years to unfold. The ability to gather and insert objective feedback into a model, to update it, and to make a better decision the next time just isn't present.

None of these caveats call into question the considerable power of decision analysis and predictive models in so many domains. They help underscore the main point: an appreciation of decision analytics is important, but an understanding of when these techniques are useful and of their limitations is essential, too.

• • •

Most executives today would probably admit that they are overwhelmed by the volume and complexity of the decisions they face and are grateful when models may relieve some of the burden. But they need to be careful. Decision models are often so impressive that it's easy to be seduced by them and to overlook the need to use them wisely. As University of Calgary associate professor Jeremy Fox observed, the growing popularity of "technically sophisticated, computationally intensive statistical approaches" has an unfortunate side effect: a "shut up and calculate the numbers" ethos, rather than one that promotes critical thinking and stimulates ideas about what the numbers actually mean.[7] Before leaders and their teams apply models, they should step back and consider their ability to influence the outcome. When it is high, the answer isn't to ignore the data and fly blind, but to establish priorities for tipping the scales through the strength and confidence that are hallmarks of effective leadership.

[7] *Oikos Online*, "Frequentist vs. Bayesian statistics: Resources to help you choose," blog entry by Jeremy Fox, October 11, 2011, oikosjournal.wordpress.com.

**Phil Rosenzweig** is a professor of strategy and international business at the International Institute for Management Development (IMD), in Lausanne, Switzerland. This article is adapted from his new book, *Left Brain, Right Stuff: How Leaders Make Winning Decisions* (PublicAffairs, January 2014).