

The Big Idea

THE MEAS SUCC

OF

Most companies use the
wrong performance metrics.
Don't be one of them.

by Michael J. Mauboussin

TRUE METERS OF BUSINESS

PHOTOGRAPHY: BRUCE PETERSON

A

ABOUT A DOZEN years ago, when I was working for a large financial services firm, one of the senior executives asked me to take on a project to better understand the company's profitability. I was in the equity division, which generated fees and commissions by catering to investment managers and sought to maximize revenues by providing high-quality research, responsive trading, and coveted initial public offerings. While we had hundreds of clients, one mutual fund company was our largest. We shuttled our researchers to visit with its analysts and portfolio managers, dedicated capital to ensure that its trades were executed smoothly, and recognized its importance in the allocation of IPOs. We were committed to keeping the 800-pound gorilla happy.

Part of my charge was to understand the division's profitability by customer. So we estimated the cost we incurred servicing each major client. The results were striking and counterintuitive: Our largest customer was among our least profitable. Indeed, customers in the middle of the pack, which didn't demand substantial resources, were more profitable than the giant we fawned over.

What happened? We made a mistake that's exceedingly common in business: We measured the wrong thing. The statistic we relied on to assess our performance—revenues—was disconnected from our overall objective of profitability. As a result, our strategic and resource allocation decisions didn't support that goal. This article will reveal how this mistake permeates businesses—probably even yours—driving poor decisions and undermining performance. And it will show you how to choose the best statistics for your business goals.

Ignoring Moneyball's Message

Moneyball, the best seller by Michael Lewis, describes how the Oakland Athletics used carefully chosen statistics to build a winning baseball team on

the cheap. The book was published nearly a decade ago, and its business implications have been thoroughly dissected. Still, the key lesson hasn't sunk in. Businesses continue to use the wrong statistics.

Before the A's adopted the methods Lewis describes, the team relied on the opinion of talent scouts, who assessed players primarily by looking at their ability to run, throw, field, hit, and hit with power. Most scouts had been around the game nearly all their lives and had developed an intuitive sense of a player's potential and of which statistics mattered most. But their measures and intuition often failed to single out players who were effective but didn't look the part. Looks might have nothing to do with the statistics that are actually important: those that reliably predict performance.

Baseball managers used to focus on a basic number—team batting average—when they talked about scoring runs. But after doing a proper statistical analysis, the A's front office recognized that a player's ability to get on base was a much better predictor of how many runs he would score. Moreover, on-base percentage was underpriced relative to other abilities in the market for talent. So the A's looked for players with high on-base percentages, paid less attention to batting averages, and discounted their gut sense. This allowed the team to recruit winning players without breaking the bank.

Many business executives seeking to create shareholder value also rely on intuition in selecting statistics. The metrics companies use most often to measure, manage, and communicate results—often called key performance indicators—include financial measures such as sales growth and earnings per share (EPS) growth in addition to nonfinancial measures such as loyalty and product quality. Yet, as we'll see, these have only a loose connection to the objective of creating value. Most executives continue to lean heavily on poorly chosen statistics, the equivalent of using batting averages to predict runs.

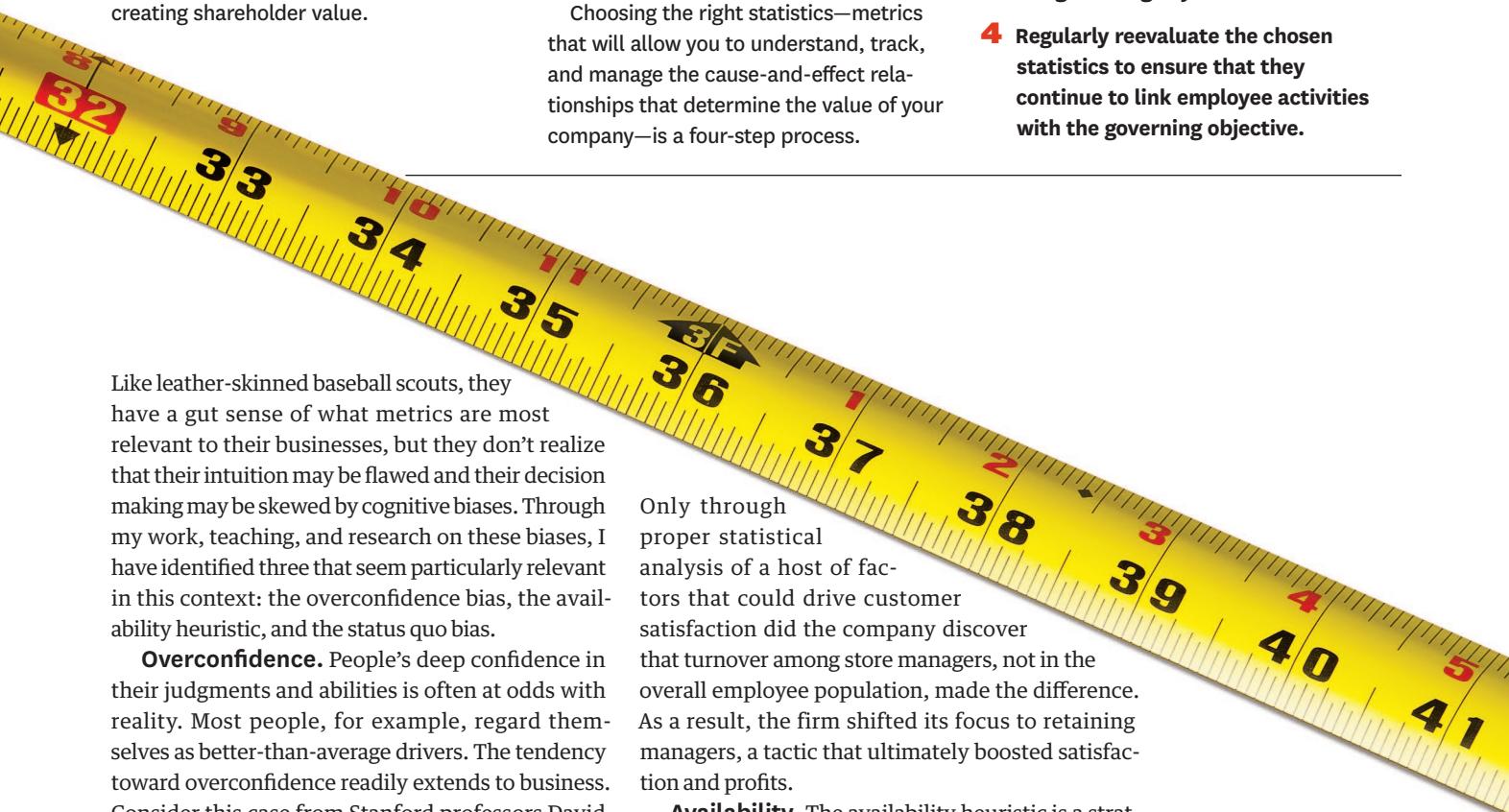
Idea in Brief

The statistics that companies use most often to track and communicate performance include financial measures such as sales and earnings per share growth as well as nonfinancial measures such as loyalty and product quality. Yet these have only a flimsy connection to the objective of creating shareholder value.

Useful statistics have two qualities. They are persistent, showing that the outcome of an action at one time will be similar to the outcome of the same action at a later time; and they are predictive, demonstrating a causal relationship between the action and the outcome being measured.

Choosing the right statistics—metrics that will allow you to understand, track, and manage the cause-and-effect relationships that determine the value of your company—is a four-step process.

- 1 Define your governing objective.**
- 2 Develop a theory of cause and effect to assess presumed drivers of value.**
- 3 Identify the specific activities that employees need to do to help achieve the governing objective.**
- 4 Regularly reevaluate the chosen statistics to ensure that they continue to link employee activities with the governing objective.**



Like leather-skinned baseball scouts, they have a gut sense of what metrics are most relevant to their businesses, but they don't realize that their intuition may be flawed and their decision making may be skewed by cognitive biases. Through my work, teaching, and research on these biases, I have identified three that seem particularly relevant in this context: the overconfidence bias, the availability heuristic, and the status quo bias.

Overconfidence. People's deep confidence in their judgments and abilities is often at odds with reality. Most people, for example, regard themselves as better-than-average drivers. The tendency toward overconfidence readily extends to business. Consider this case from Stanford professors David Larcker and Brian Tayan: The managers of a fast-food chain, recognizing that customer satisfaction was important to profitability, believed that low employee turnover would keep customers happy. "We just know this is the key driver," one executive explained. Confident in their intuition, the executives focused on reducing turnover as a way to improve customer satisfaction and, presumably, profitability.

As the turnover data rolled in, the executives were surprised to discover that they were wrong: Some stores with high turnover were extremely profitable, while others with low turnover struggled.

Only through proper statistical analysis of a host of factors that could drive customer satisfaction did the company discover that turnover among store managers, not in the overall employee population, made the difference. As a result, the firm shifted its focus to retaining managers, a tactic that ultimately boosted satisfaction and profits.

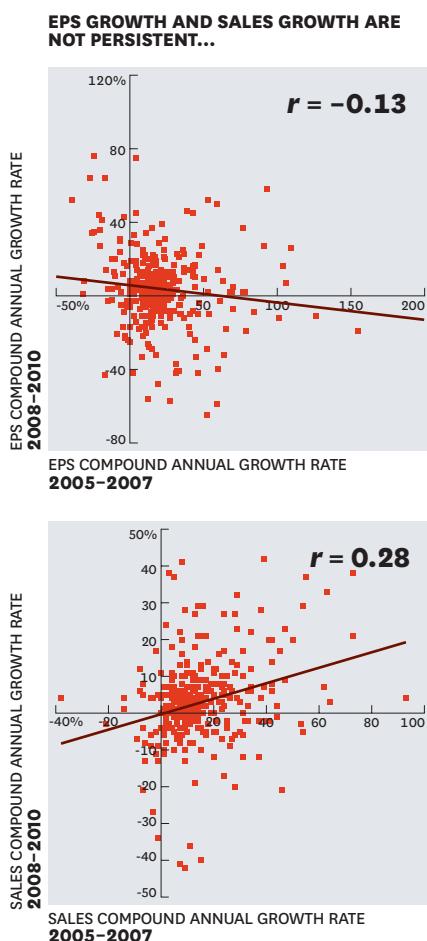
Availability. The availability heuristic is a strategy we use to assess the cause or probability of an event on the basis of how readily similar examples come to mind—that is, how "available" they are to us. One consequence is that we tend to overestimate the importance of information that we've encountered recently, that is frequently repeated, or that is top of mind for other reasons. For example, executives generally believe that EPS is the most important measure of value creation in large part because of vivid examples of companies whose stock rose after they exceeded EPS estimates or fell abruptly after coming up short. To many executives, earnings

People's deep confidence in their judgments and abilities is often at odds with reality.

The Problem with Popular Measures

The most useful statistics are *persistent* (they show that the outcome of an action at one time will be similar to the outcome of the same action at another time) and *predictive* (they link cause and effect, predicting the outcome being measured). Statisticians assess a measure's persistence and its predictive value by examining the coefficient of correlation: the degree of the linear relationship between variables in a pair of distributions. Put simply, if there is a strong relationship between two sets of variables (say a group of companies' sales growth in two different periods), plotting the points on a graph like the ones shown here produces a straight line. If there's no relationship between the variables, the points will appear to be randomly scattered, in this case showing that sales growth in the first period does not predict sales growth in the second.

In comparing the variable "sales growth" in two periods, the coefficient of correlation, r , falls in the range of 1.00 to -1.00. If each company's sales growth is the same in both periods (a



perfect positive correlation), $r = 1.00$ —a straight line. (The values need not be equal to produce a perfect correlation; any straight line will do.) If sales growth in the two periods is unrelated (there is zero correlation), $r = 0$ —a random pattern. If increases in one period match decreases in the other (a perfect inverse correlation), $r = -1.00$ —also a straight line. Even a quick glance can tell you whether there is a high correlation between the variables (the points are tightly clustered and linear) or a low correlation (they're randomly scattered).

The closer to 1.00 or -1.00 the coefficient of correlation is, the more persistent and predictive the statistic. The closer to zero, the less persistent and predictive the statistic.

Let's examine the persistence of two popular measures: EPS growth and sales growth.

The figures to the left show the coefficient of correlation for EPS growth and sales growth for more than 300 large nonfinancial companies in the

growth seems like a reliable cause of stock-price increases because there seems to be so much evidence to that effect. But, as we'll see, the availability heuristic often leads to flawed intuition.

Status quo. Finally, executives (like most people) would rather stay the course than face the risks that come with change. The status quo bias derives in part from our well-documented tendency to avoid a loss even if we could achieve a big gain. A business consequence of this bias is that even when performance drivers change—as they invariably do—executives often resist abandoning existing metrics in favor of more-suitable ones. Take the case of a subscription business such as a wireless telephone provider. For a new entrant to the market, the acquisition rate of new customers is the most important performance metric. But as the company matures, its emphasis should probably shift from adding customers to better managing the ones it has by, for instance, selling them additional services or reducing churn. The pull of the status quo, however, can

inhibit such a shift, and so executives end up managing the business with stale statistics.

Considering Cause and Effect

To determine which statistics are useful, you must ask two basic questions. First, what is your objective? In sports, it is to win games. In business, it's usually to increase shareholder value. Second, what factors will help you achieve that objective? If your goal is to increase shareholder value, which activities lead to that outcome?

What you're after, then, are statistics that reliably reveal cause and effect. These have two defining characteristics: They are *persistent*, showing that the outcome of a given action at one time will be similar to the outcome of the same action at another time; and they are *predictive*—that is, there is a causal relationship between the action the statistic measures and the desired outcome.

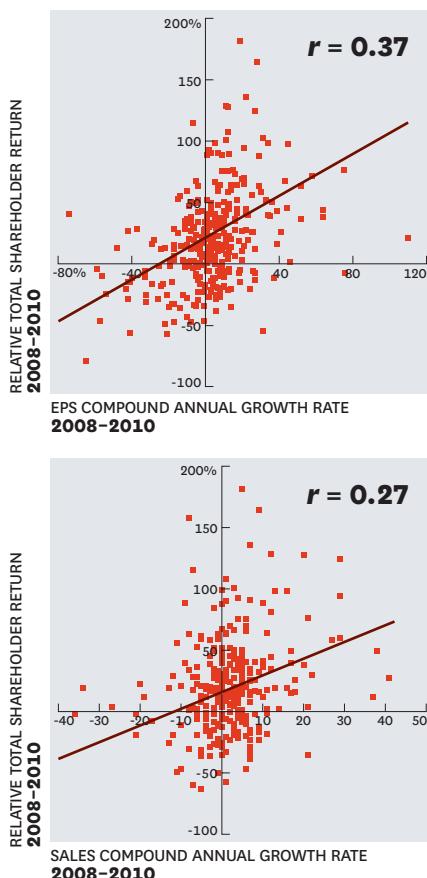
Statistics that assess activities requiring skill are persistent. For example, if you measured the perfor-

United States. The compounded annual growth rates from 2005 to 2007, on the horizontal axes, are compared with the rates from 2008 to 2010, on the vertical axes. If EPS and sales growth were highly persistent and, therefore, dependent on factors the company could control, the points would cluster tightly on a straight line. But in fact they're widely scattered, revealing the important role of chance or luck. The correlation is negative and relatively weak ($r = -0.13$) for EPS growth but somewhat higher ($r = 0.28$) for sales growth. This is consistent with the results of large-scale studies.

Next, we'll look at the predictive value of EPS growth and sales growth by examining the correlation of each with shareholder returns.

In the figures to the right, adjusted EPS growth and sales growth are on the horizontal axes. The vertical axes are the total return to shareholders for each company's stock less the total return for the S&P 500. Adjusted EPS

...AND THEY DON'T DO A GOOD JOB OF PREDICTING VALUE CREATION.



mance of a trained sprinter running 100 meters on two consecutive days, you would expect to see similar times. Persistent statistics reflect performance that an individual or organization can reliably control through the application of skill, and so they expose causal relationships.

It's important to distinguish between skill and luck. Think of persistence as occurring on a continuum. At one extreme the outcome being measured is the product of pure skill, as it was with the sprinter, and is very persistent. At the other, it is due to luck, so persistence is low. When you spin a roulette wheel, the outcomes are random; what happens on the first spin provides no clue about what will happen on the next.

To be useful, statistics must also predict the result you're seeking. Recall the Oakland A's recognition that on-base percentage told more about a player's likelihood of scoring runs than his batting average did. The former statistic reliably links a cause (the ability to get on base) with an effect (scoring runs). It

is also more persistent than batting average because it incorporates more factors—including the ability to get walked—that reflect skill. So we can conclude that a team's on-base percentage is better for predicting the performance of a team's offense.

All this seems like common sense, right? Yet companies often rely on statistics that are neither very persistent nor predictive. Because these widely used metrics do not reveal cause and effect, they have little bearing on strategy or even on the broader goal of earning a sufficient return on investment.

Consider this: Most corporations seek to maximize the value of their shares over the long term. Practically speaking, this means that every dollar a company invests should generate more than one dollar in value. What statistics, then, should executives use to guide them in this value creation? As we've noted, EPS is the most popular. A survey of executive compensation by Frederic W. Cook & Company found that it is the most popular measure of corporate performance, used by about half

growth shows a reasonably good correlation with increasing shareholder value ($r = 0.37$), so it is somewhat predictive. The problem is that forecasting earnings is difficult because, as we saw in the previous analysis, EPS growth in one period tells you little about what will happen in another. Earnings data may be moderately predictive of shareholder returns, but they are not persistent.

Using sales growth as a gauge of value creation falls short for a different reason. While sales growth is more persistent than EPS growth, it is less strongly correlated with relative total returns to shareholders ($r = 0.27$). In other words, sales-growth statistics may be somewhat persistent, but they're not very predictive.

Thus the two most popular measures of performance have limited value in predicting shareholder returns because neither is both persistent *and* predictive.

The Perils of Intuition

To identify useful statistics, you must have a solid grasp of cause and effect. If you don't understand the sources of customer satisfaction, for example, you can't identify the metrics that will help you improve it. This seems obvious, but it's surprising how often people assign the wrong cause to an outcome. This failure results from an innate desire to find cause and effect in every situation—to create a narrative that explains how events are linked even when they're not.

Consider this: The most common method for teaching business management is to find successful businesses, identify their common practices, and recommend that managers imitate them. Perhaps the best-known book using this method is Jim Collins's *Good to Great*. Collins and his team analyzed thousands of companies and isolated 11 whose performance went from good to great. They then identified the practices that they believed had caused those companies to improve—including leadership, people, a fact-based approach, focus, discipline, and the use of technology—and suggested that other companies adopt them to achieve the same great results. This formula is intuitive, includes some compelling narrative, and has sold millions of books.

If causality were clear, this approach would work. The trouble is that the performance of a company almost always depends on both skill and luck, which means that a given strategy will succeed only part of the time. Some companies using the strategy will succeed; others will fail. So attributing a firm's success to a specific strategy may be wrong if you sample only the winners.

The more important question is, How many of the companies that tried the strategy actually succeeded?

Jerker Denrell, a professor of strategy at Oxford, calls this the “undersampling of failure.” He argues that because firms with poor performance are unlikely to survive, they are absent from the group under observation. Say two companies pursue the same strategy, and one succeeds because of luck while the other fails. Since we draw our sample from the outcome, not the strategy, we observe the successful company and assume that the favorable outcome was the result of skill and overlook the influence of luck. We connect cause and effect where there is no connection.

The lesson is clear: When luck plays a part in determining the consequences of your actions—as is often the case in business—you don't want to study success to identify good strategy but rather study strategy to see whether it consistently led to success. Statistics that are persistent and predictive, and so reliably link cause and effect, are indispensable in that process.



of all companies. Researchers at Stanford Graduate School of Business came to the same conclusion. And a survey of 400 financial executives by finance professors John Graham, Campbell Harvey, and Shiva Rajgopal found that nearly two-thirds of companies placed EPS first in a ranking of the most important performance measures reported to outsiders. Sales revenue and sales growth also rated highly for measuring performance and for communicating externally.

But will EPS growth actually create value for shareholders? Not necessarily. Earnings growth and value creation can coincide, but it is also possible to increase EPS while destroying value. EPS growth is good for a company that earns high returns on invested capital, neutral for a company with returns equal to the cost of capital, and bad for companies with returns below the cost of capital. Despite this, many companies slavishly seek to deliver EPS growth, even at the expense of value creation. The

survey by Graham and his colleagues found that the majority of companies were willing to sacrifice long-term economic value in order to deliver short-term earnings. Theory and empirical research tell us that the causal relationship between EPS growth and value creation is tenuous at best. Similar research reveals that sales growth also has a shaky connection to shareholder value. (For a detailed examination of the relationship between earnings growth, sales growth, and value, see the exhibit “The Problem with Popular Measures.”)

Of course, companies also use nonfinancial performance measures, such as product quality, workplace safety, customer loyalty, employee satisfaction, and a customer's willingness to promote a product. In their 2003 HBR article, accounting professors Christopher Ittner and David Larcker wrote that “most companies have made little attempt to identify areas of nonfinancial performance that might advance their chosen strategy. Nor have they

demonstrated a cause-and-effect link between improvements in those nonfinancial areas and in cash flow, profit, or stock price.” The authors’ survey of 157 companies showed that only 23% had done extensive modeling to determine the causes of the effects they were measuring. The researchers suggest that at least 70% of the companies they surveyed didn’t consider a nonfinancial measure’s persistence or its predictive value. Nearly a decade later, most companies still fail to link cause and effect in their choice of nonfinancial statistics.

But the news is not all bad. Ittner and Larcker did find that companies that bothered to measure a nonfinancial factor—and to verify that it had some real effect—earned returns on equity that were about 1.5 times greater than those of companies that didn’t take those steps. Just as the fast-food chain boosted its performance by determining that its key metric was store manager turnover, not overall employee turnover, companies that make proper links between nonfinancial measures and value creation stand a better chance of improving results.

Picking Statistics

The following is a process for choosing metrics that allow you to understand, track, and manage the cause-and-effect relationships that determine your company’s performance. I will illustrate the process in a simplified way using a retail bank that is based on an analysis of 115 banks by Venky Nagar of the University of Michigan and Madhav Rajan of Stan-

ford. Leave aside, for the moment, which metrics you currently use or which ones Wall Street analysts or bankers say you should. Start with a blank slate and work through these four steps in sequence.

1 Define your governing objective. A clear objective is essential to business success because it guides the allocation of capital. Creating economic value is a logical governing objective for a company that operates in a free market system. Companies may choose a different objective, such as maximizing the firm’s longevity. We will assume that the retail bank seeks to create economic value.

2 Develop a theory of cause and effect to assess presumed drivers of the objective. The three commonly cited financial drivers of value creation are sales, costs, and investments. More-specific financial drivers vary among companies and can include earnings growth, cash flow growth, and return on invested capital.

Naturally, financial metrics can’t capture all value-creating activities. You also need to assess nonfinancial measures such as customer loyalty, customer satisfaction, and product quality, and determine if they can be directly linked to the financial measures that ultimately deliver value. As we’ve discussed, the link between value creation and financial

CONTINUED ON PAGE 56

Companies that link nonfinancial measures and value creation stand a better chance of improving results.



and nonfinancial measures like these is variable and must be evaluated on a case-by-case basis.

In our example, the bank starts with the theory that customer satisfaction drives the use of bank services and that usage is the main driver of value. This theory links a nonfinancial and a financial driver. The bank then measures the correlations statistically to see if the theory is correct and determines that satisfied customers indeed use more services, allowing the bank to generate cash earnings growth and attractive returns on assets, both indicators of value creation. Having determined that customer satisfaction is persistently and predictively linked to returns on assets, the bank must now figure out which employee activities drive satisfaction.

3 Identify the specific activities that employees can do to help achieve the governing objective.

The goal is to make the link between your objective and the measures that employees can control through the application of skill. The relationship between these activities and the objective must also be persistent and predictive.

In the previous step, the bank determined that customer satisfaction drives value (it is predictive). The bank now has to find reliable drivers of customer satisfaction. Statistical analysis shows that the rates consumers receive on their loans, the speed of loan processing, and low teller turnover all affect customer satisfaction. Because these are within the control of employees and management, they are

persistent. The bank can use this information to, for example, make sure that its process for reviewing and approving loans is quick and efficient.

4 Evaluate your statistics. Finally, you must regularly reevaluate the measures you are using to link employee activities with the governing objective. The drivers of value change over time, and so must your statistics. For example, the demographics of the retail bank's customer base are changing, so the bank needs to review the drivers of customer satisfaction. As the customer base becomes younger and more digitally savvy, teller turnover becomes less relevant and the bank's online interface and customer service become more so.

COMPANIES HAVE access to a growing torrent of statistics that could improve their performance, but executives still cling to old-fashioned and often flawed methods for choosing metrics. In the past, companies could get away with going on gut and ignoring the right statistics because that's what everyone else was doing. Today, using them is necessary to compete. More to the point, identifying and exploiting them before rivals do will be the key to seizing advantage. □

HBR Reprint R1210B

 **Michael J. Mauboussin** is the chief investment strategist at Legg Mason Capital Management and an adjunct professor of finance at Columbia Business School. He is the author of *The Success Equation* (Harvard Business Review Press, forthcoming), from which this article was developed.



"I hear they call you the hard-copy guy."

Harvard Business Review Notice of Use Restrictions, May 2009

Harvard Business Review and Harvard Business Publishing Newsletter content on EBSCOhost is licensed for the private individual use of authorized EBSCOhost users. It is not intended for use as assigned course material in academic institutions nor as corporate learning or training materials in businesses. Academic licensees may not use this content in electronic reserves, electronic course packs, persistent linking from syllabi or by any other means of incorporating the content into course resources. Business licensees may not host this content on learning management systems or use persistent linking or other means to incorporate the content into learning management systems. Harvard Business Publishing will be pleased to grant permission to make this content available through such means. For rates and permission, contact permissions@harvardbusiness.org.