A theory of relativity
Setting priorities and goals for financial performance improvement

BY MICHAEL E. RAYNOR, MUMTAZ AHMED, DEREK PANKRATZ, AND ROB DEL VICARIO
> ILLUSTRATION BY THOMAS KUHLENBECK
Most every budget or strategic plan includes financial performance goals of some sort. In many cases, whether or not a company meets its targets for growth or profitability is the yardstick for measuring success or failure, the basis upon which financial incentives are paid out to managers, and a key driver of shareholder returns.

The specific financial objectives companies set for themselves can matter a great deal. Setting targets that are too aggressive can mean that even the best efforts go unrewarded, leaving people demoralized.\(^1\) Worse, in an effort to win the rewards that come with achieving performance goals, people can end up motivated to cut corners or to resort to unethical or illegal behavior that they would otherwise loathe even to contemplate.\(^2\) Finally, setting the right goal for the wrong objective—say, focusing on increasing sales when the problem is depressed profitability—can leave an organization struggling even when its goals have been met.\(^3\)

In our view, setting the right goals and priorities involves at least three questions. First, every leadership team must answer “How are we doing?” Setting financial performance targets is very often informed by benchmarking of some sort: an assessment of a company’s performance relative to some peer group. How much and fast a company needs to improve is, quite reasonably, a function of how well it is doing compared with the overall economy, its industry, or its nearest competitors. The answer to this question leads naturally to the second question: “What should we improve?” Should the emphasis be on profitability? Growth? Something else?

Having assessed one’s relative position and set performance improvement priorities, the third question is “By how much do we need to improve?” In light of what the company’s performance is, what should it be, and how long should it take to get there?
relative

performance improvement
Unfortunately, our research suggests that many managers answer these questions using assessments of relative performance that are based on incomplete facts, misinformed intuition, and inappropriate analysis. Worse, these errors can bias in highly dysfunctional ways the ability of managers to interpret the nature of the opportunities a company faces and its probability of success.

There is a way to measure a company’s relative performance, set targets, and estimate the probability of achieving specified targets over different time periods that allows managers to resist the pull of methods that are often widely used, yet can be dangerously misleading. Including this approach as part of a company’s goal-setting process can lead to more realistic targets.

**OF PERFORMANCE BOTH ABSOLUTE AND RELATIVE**

Knowing that a company is profitable or losing money, or that it is growing or shrinking, is of course revealing. These assessments of corporate performance make no reference to other companies’ results, and so we think of them as absolute. It is difficult to set performance targets based only on absolute measures, however. In general terms, higher profitability and stronger growth are better, but this “Buzz Lightyear approach” to target setting (“to infinity . . . and beyond”) hardly seems satisfactory.

To develop an aspirational but credible goal, we typically turn to a relative assessment of past and current performance. We compare a company’s results with those of a relevant peer group and set targets for improvement that translate into desired increases in relative rank. This is the thinking behind benchmarking: One seeks to increase return on assets (ROA) from 8 to 10 percent, not because of the significance of the extra two percentage points of ROA but because that increase might mean moving from the median to the top quartile in one’s industry. In other words, absolute performance sets a performance floor, but relative performance tells us where the ceiling is.

Unfortunately, there appears to be no generally accepted, objective, quantitative method for measuring relative rank. Worse, the methods often used to identify a peer group tend to provide wildly misleading results for one of two reasons: A comparison set is too large and diverse, or it is too small and homogenous.

**The telescope problem**

One way to evaluate relative performance is to compare a broad cross-section of companies and construct a simple rank ordering. Such approaches are quite common and popular, and are a staple of popular business publications. We might call this “benchmarking against the market.”
The problem, of course, is that such a wide lens fails to consider the context in which a company finds itself. Stars seen through a hobbyist’s telescope might appear to be of similar sizes and distances from Earth when in fact they are orders-of-magnitude different in surface area and light-years apart; some may have already ceased to exist by the time their light reaches our eyes. Similarly, can a company operating in a low-margin sector reasonably hope to achieve the performance of another enjoying the rising tide of a structurally more profitable industry? Should a $50 billion company set a growth target based on benchmarks set by companies working off a base of $50 million?

The microscope problem

Knowing how the overall economy is doing or how well all publicly traded companies are performing might be useful background, but we expect that most managers go beyond this gross benchmarking and develop a more carefully chosen peer group so that relevant relative performance targets can be set. A survey from 2012 indicates that roughly 75 percent of executives around the world use this form of benchmarking, a proportion largely unchanged in almost a decade.

Specific applications vary, but managers often begin by restricting their benchmarks by industry, sector, and size. The performances of the companies in this relatively homogenous group are then compared over a specified time period—say, three or five years. This yields a series of rankings, which become management’s assessment of what their own company’s performance really means.

Unfortunately, optimizing for comparability too often creates its own problems: Such a narrow basis of comparison undermines the ability to make reliable inferences about a company’s true relative position because there are too few comparisons to make. You may find yourself the top performer among your four closest peers, earning you bragging rights and relief that you’re not last. But should you take comfort in this result? After all, there’s a 25 percent chance of finding yourself atop the group through chance alone.

Worse, small groups are more susceptible to extreme outcomes and higher fluctuations, both positive and negative, than large ones. Consequently, it is much more difficult to identify the signal in the noise when the number of peers is low. It is too easy to mistake what is actually an extended streak of good (or bad) luck for true breakthrough (or fall-through) performance by your competitors and, as a result, to end up striving for the unattainable or resting on false laurels.

Anchors aweigh

If we could accurately and completely assess addressable opportunities and the capabilities an organization can use to pursue them, accurate relative ranking
A THEORY OF RELATIVITY

might not matter so much. Corporate planners have long sought to understand a company’s potential independently of its relative performance, focusing instead on each company’s relative competitive position and organizational strengths or weaknesses.6 Ironically, it is precisely because such analysis is critical to setting meaningful goals that benchmarks are so crucial. The evaluation of capabilities and opportunities is unavoidably subjective, and so their implications for performance targets are at least as much imposed as inferred.

In this context, performance benchmarks become anchors that shade and color our interpretation of very complex and ambiguous data. A well-established body of research in psychology and behavioral economics tells us that anchors of this sort—the starting point for generating an estimate of an unknown quantity—insensibly, but consistently and dramatically, influence our ultimate choices.7

For example, in a classic demonstration, subjects were assigned a random number between 0 and 100 generated by the spin of a wheel. They were then asked to estimate the percentage of African countries in the United Nations. The random number they were assigned had a dramatic impact on their estimate. For example, the median estimate of those who received 10 as their anchor was 25 percent. The median for the group that received 65 as an anchor was 45 percent—a 20 percentage point difference, despite the fact that the participants knew their anchor was irrelevant and randomly assigned.8 And so powerful is this bias that clearly implausible anchors can skew results, even when subjects are alerted to the potential impact.9

THE STRATEGIC IMPORTANCE OF FINANCIAL BENCHMARKING

Keeping tabs only on companies sufficiently “like” yours is an increasingly risky approach to viewing your competitive environment. While phrases such as “disruption” and “unprecedented competitive pressure” are bandied about with abandon, threats can emerge from nontraditional corners of the economy.

Herein lies another limitation of classic benchmarking. By narrowly defining the competitive environment, it becomes all too easy to miss these emerging threats. Even as you track the same five or seven peers year after year—each year concluding you’re near the top—your business is being eaten away by a new competitor. Ironically, the conclusion that you lead the pack may not be wrong, since your traditional competitors’ businesses are also being eroded! Without constant vigilance and an expansive definition of what constitutes a “competitor,” you leave yourself vulnerable. Increasingly, we want to compare apples to oranges . . . and to bananas, and anything else that might be sprouting in the undergrowth. To do that, we need a better approach to benchmarking, one that includes very different companies but allows for valid comparisons despite those differences.
When it comes to setting performance goals, the anchor in our decision making is our assessment of a company’s current relative position. How well we think a company is doing today will influence both our perceived need for improvement and how we interpret its prospects for improvement. If our benchmark places a company in the bottom quartile, we may be biased toward seeing opportunities to move up; if we think a company is besting relevant rivals, it might be more difficult to identify attractive white spaces and easier to ignore potential threats.

In short, we cannot avoid anchoring, but, as we will demonstrate below, some of the anchors used are misleading.

HOW ARE WE DOING?

Our first challenge, then, is to develop a method that can answer the “How are we doing?” question but that is not subject to the “telescope” and “microscope” problems. We want to take full advantage of the sizable quantity of company data at our disposal, but we also want to take into account the specific circumstances of each company.

Our approach relies on a combination of semiparametric statistical techniques and simulations. We use quantile regression models to strip the effects of industry, size, and year from each company’s financial performance. Because these adjustments are based on a population-level regression, each company’s rank is compared with the full population of all other US-based public companies. Just as a

Figure 1. Comparing ROA and revenue growth percentile ranks using different methods

A. ROA
handicap allows golfers of different abilities to play on even terms, so our modeling approach enables us to compare companies facing drastically different opportunities and constraints.

We also want to characterize a company’s performance at a point in time in the context of its performance over time. To avoid being fooled by single-year aberrations, we create a dynamic moving average, more heavily weighing performance closest to the focal year. This attenuates the often-drastic year-over-year fluctuations in performance that can be driven by anything from a merger to a one-time write-down or asset sale. Finally, rather than picking an arbitrary timeframe like three or five years to look at a company’s performance, the time period over which the moving average is calculated is inferred from the volatility of the underlying financial measure.\textsuperscript{11}

Such a rigorous and complex method is only justified if the results are materially different from what a simpler approach would yield. Consider a company like FeCo, a real but anonymized firm that manufactures metal goods. In 2013, FeCo saw revenue contract over 16 percent in real terms. When viewed through the telescope and ranked against the roughly 5,000 active US-based public companies in the same year, FeCo is in the 12th percentile, worse than nearly 90 percent of all companies. Yet, looking through the microscope and compared with its closest peers in the same industry and of roughly the same size, FeCo’s five-year average growth places it at No. 1 out of 3. So perhaps all is well.
The story changes when we apply our approach. FeCo’s long-run weighted average percentile rank for revenue growth is 46.9, solidly in the middle of the pack. By attenuating the extremes of the “telescope” and “microscope” approaches, we can arrive at a truer picture of the underlying reality. In this case, FeCo’s performance is neither quite so dire nor quite as rosy as simpler approaches to benchmarking would suggest.

FeCo’s story is not unique. Figure 1 plots the ROA and revenue growth percentile ranks for the “telescope” and “microscope” approaches against our method.12

The red lines in the figure represent simple fitted linear regression lines, and they suggest a very weak relationship subject to significant variation.13 The average absolute difference in percentile ranks between our method and the more common approaches described earlier is between 18 and 25 percentile ranks across profitability and growth. That means, on average, a company might consider itself to be in the top quarter of its peer group, but really it could be no better than middle of the road.

Worse, there are many hundreds of companies in the upper-left and lower-right quadrants of these charts. It is unlikely that savvy managers would believe their companies to be first when in fact they are last, but by anchoring on such a misleading benchmark, the entire goal-setting process could be derailed.

Applying our method yields insights that are not readily intuited, even by those well positioned to understand a company’s absolute and relative performance. We
surveyed 301 executives from large US-based companies, asking them to report their absolute performance (an ROA of 5 percent, for example). We also asked them to estimate what that performance was as a relative percentile rank, taking into account their company’s industry and size. We then used our statistical model to translate their reported absolute performance level into a percentile rank, adjusting for industry and size, and compared their self-reported estimates with our results.

The “bee swarm” pattern in figure 2 suggests that few were able to translate their absolute performance into relative terms with any accuracy. Indeed, the correlation between the two estimates for profitability and growth measures was just 0.13, and the median absolute error exceeded 20 percentiles. That suggests, again, that a company may be solidly mediocre yet believe it is in the top quartile of performers. Or it may perceive itself as falling behind when it is no worse than average.

These results closely parallel two earlier survey efforts we undertook. With over 800 executives surveyed overall, we have seen little evidence that business leaders have a strong grasp of their relative position. These results suggest that relying on intuition to estimate relative performance, as a rule, is subject to sizable estimation errors, with radical differences all too common and no dominant direction of bias.

Of course, we cannot claim to have “the answer,” since no single approach can begin to address the myriad complexities and idiosyncrasies of particular companies’ circumstances. But our approach offers a more robust, quantitative starting point for discussions of performance, priorities, and goals. And without a well-founded
understanding of “How are we doing?” we can only guess at the answers to our next two questions: What should we improve, and by how much?

WHAT SHOULD WE IMPROVE?

A company can attain financial success on multiple possible dimensions—profitability and growth, for example. How should priorities be set, and how can we avoid focusing on the wrong areas? Our answer lies at the intersection of relative and absolute performance, and is summarized in figure 3. Start with companies in the northwest quadrant. They are doing well enough in absolute terms in that they are solvent or growing. Many will appear to be doing quite well, perhaps with double-digit ROA or growth numbers.

The natural conclusion might be that there’s no room left to improve, and if performance priorities are anchored on those absolute figures, attention and resources are likely to shift elsewhere. But when we augment our performance picture with relative standings, it becomes clear that these companies are leaving money on the table. Given their circumstances, even greater heights are possible.
A THEORY OF RELATIVITY

The challenge may be even greater for companies with the opposite performance profile (lower-right quadrant). Faced with flat or declining profitability or growth, the seemingly irresistible temptation is to focus on those measures in the belief that they have the greatest need or greatest potential for improvement. Our analysis, however, suggests these companies are already near the upper limit of what is feasible, given the structural constraints they face.

For example, for the largest companies in slow-growing industries, low or even negative growth might place them in the top quartile of relative performance. Here performance improvement efforts risk mimicking Sisyphus, pushing his boulder uphill only to have it roll back down and being forced to start again. The relative performance analysis suggests that if companies want to see significant gains in absolute terms, they had best look outside their traditional businesses.

The situations for companies in the remaining two quadrants are more straightforward. For those with low absolute and relative performance, the message is clear: all hands on deck. There is a need to improve in absolute terms as well as a need for the requisite headroom to achieve that improvement. Among those in the enviable
position of having high performance in both absolute and relative terms, the challenge is to stay the course. This requires both vigilance against complacency and the courage to resist the urge to “climb past the summit.” At the highest levels of
performance, dramatic improvements are unlikely or even mathematically impossible. Major initiatives to boost profitability or expand revenue are likely to fall short of expectations and could prove to be dangerous distractions from the important work of sustaining already-high levels of performance.

**BY HOW MUCH DO WE NEED TO IMPROVE?**

Knowing that a company is in the 63rd percentile says nothing about whether or how much its performance should improve. Depending on a company’s circumstances, aggressive targets or conservative ones can make perfect sense. But we should insist on going in with our eyes open, with as comprehensive an assessment of the probability of success as is feasible. An extension of our method allows us to answer our third question “By how much?” and so anchor specific goals in similarly objective analysis.

One way to think about the likelihood of hitting a performance target is to consider how frequently other companies have made similar improvements. Using over four decades of data on US-based public companies, we constructed a 100 x 100 percentile transition probability matrix that captures the frequency with which companies have moved from one percentile rank to another in one year on a given performance measure.

For example, all else equal, the probability that a company will improve from the 60th percentile of revenue growth to the 65th or better is about 0.38. In contrast, the

**Figure 4. ROA decile transition probability matrix**

<table>
<thead>
<tr>
<th>Starting decile</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.643</td>
<td>0.189</td>
<td>0.050</td>
<td>0.023</td>
<td>0.013</td>
<td>0.007</td>
<td>0.007</td>
<td>0.005</td>
<td>0.004</td>
<td>0.059</td>
</tr>
<tr>
<td>1</td>
<td>0.217</td>
<td>0.366</td>
<td>0.172</td>
<td>0.078</td>
<td>0.044</td>
<td>0.029</td>
<td>0.021</td>
<td>0.017</td>
<td>0.008</td>
<td>0.048</td>
</tr>
<tr>
<td>2</td>
<td>0.064</td>
<td>0.203</td>
<td>0.306</td>
<td>0.185</td>
<td>0.091</td>
<td>0.052</td>
<td>0.032</td>
<td>0.021</td>
<td>0.009</td>
<td>0.036</td>
</tr>
<tr>
<td>3</td>
<td>0.028</td>
<td>0.099</td>
<td>0.199</td>
<td>0.298</td>
<td>0.188</td>
<td>0.084</td>
<td>0.048</td>
<td>0.022</td>
<td>0.009</td>
<td>0.026</td>
</tr>
<tr>
<td>4</td>
<td>0.015</td>
<td>0.053</td>
<td>0.107</td>
<td>0.183</td>
<td>0.296</td>
<td>0.193</td>
<td>0.080</td>
<td>0.033</td>
<td>0.012</td>
<td>0.028</td>
</tr>
<tr>
<td>5</td>
<td>0.009</td>
<td>0.031</td>
<td>0.061</td>
<td>0.100</td>
<td>0.181</td>
<td>0.314</td>
<td>0.196</td>
<td>0.061</td>
<td>0.019</td>
<td>0.028</td>
</tr>
<tr>
<td>6</td>
<td>0.006</td>
<td>0.021</td>
<td>0.037</td>
<td>0.061</td>
<td>0.091</td>
<td>0.178</td>
<td>0.347</td>
<td>0.178</td>
<td>0.038</td>
<td>0.044</td>
</tr>
<tr>
<td>7</td>
<td>0.004</td>
<td>0.013</td>
<td>0.023</td>
<td>0.030</td>
<td>0.040</td>
<td>0.071</td>
<td>0.164</td>
<td>0.425</td>
<td>0.156</td>
<td>0.073</td>
</tr>
<tr>
<td>8</td>
<td>0.002</td>
<td>0.006</td>
<td>0.011</td>
<td>0.013</td>
<td>0.017</td>
<td>0.026</td>
<td>0.045</td>
<td>0.144</td>
<td>0.611</td>
<td>0.126</td>
</tr>
<tr>
<td>9</td>
<td>0.027</td>
<td>0.034</td>
<td>0.041</td>
<td>0.035</td>
<td>0.036</td>
<td>0.041</td>
<td>0.056</td>
<td>0.087</td>
<td>0.124</td>
<td>0.519</td>
</tr>
</tbody>
</table>

Graphic: Deloitte University Press  | DUPress.com
probability of a company improving from the 60th to the 90th percentile or above is just 0.06. Figure 4 shows an abbreviated version of the transition matrix for ROA that aggregates performance into deciles.

Of course, this does not capture the likelihood of success for a specific company. Instead, in precisely the same way that our assessment of relative performance is a sound anchor for an examination of a company’s imperatives and priorities for improvement, this assessment of the probability of success is a sound anchor for an examination of a company’s improvement strategies.

For example, if management determines that a dramatic improvement, one that has a low expected likelihood of success, is called for, then management should be prepared to pursue a more aggressive strategy. Expecting low-likelihood increases in profitability when the plan calls for little more than garden-variety efficiency improvements implies a potentially serious mismatch. On the other hand, envisioning otherwise improbable increases in growth arising from a breakthrough disruption is rather more plausible. These extreme examples might seem obvious, but the picture of corporate goal setting that emerges from our survey results is not promising.

Figure 5 displays the distributions of the estimated probabilities of meeting or exceeding ROA and growth performance targets, broken down by respondents’ self-reported estimates of how likely it is that their company will achieve that target. If respondents’ beliefs tracked the underlying likelihood of success, we would expect to see the central tendencies of the boxplots move higher on the y-axis as we move from left to right. What we see instead is almost no difference. Those who were more confident—who thought there was a 75 percent chance or better of success—have not in fact set more attainable goals. In individual cases, the optimist might be right, of course; some of these companies may hit their low-probability targets. But overall, across the sample, there is a worrying disconnect between expectations and how American companies have historically performed. Note also that almost none of our survey respondents thought their goal was very unlikely (less than 10 percent chance of success).

In short, with so little correspondence between reported chances of success and the probability of success as estimated by our method, there is too high a likelihood that the plans supporting companies’ objectives are similarly out of alignment. None of which is to say that companies should not set ambitious goals, or conservative goals for that matter. But the aggressiveness of those goals should be in line with the aggressive of one’s strategy, appetite for risk, and ability to manage that risk.
Figure 5. Estimated probability of exceeding performance goal, by respondents’ reported likelihood of success

A. ROA

Estimated probability of meeting or exceeding goal

Reported chances of success

B. Growth

Estimated probability of meeting or exceeding goal

Reported chances of success

Graphic: Deloitte University Press | DUPress.com
BETTER BENCHMARKING FOR A “BETTER” BIAS

If we were omniscient, traditional benchmarking would be obsolete. Free from the baggage of the past, we would base performance priorities and goals solely on a forward-looking assessment of the capabilities, resources, and strategy of the organization and its competitive context.

Unfortunately, our own past and sense of how we compare with others are inescapable anchors, affecting how we interpret the world around us and the goals we set for ourselves. Worse, common methods for making these comparisons are both limited and misleading. Simple rankings against all companies fail to adjust for critical context, such as the effects of industry and size. Classic “most-similar” benchmarking can create an unnecessarily small comparison group, making it difficult to distinguish the signal from the noise. Our intuitions easily lead us astray. We can end up dramatically over- or underestimating how we are doing, which can lead to misplaced priorities and unrealistic expectations for the future.

Setting the “right” targets will never be an automated process—not least because what is “right” will depend on a company’s appetite for risk, the resources at its disposal, and its competitive context. Circumstances will always matter. But since we must be biased, let us be biased as much as possible toward the underlying economic reality. Employing a rigorous, quantitative approach to performance benchmarking can serve as a better anchor around which to center discussions of how a company is doing, what it should improve, and by how much. DR

Michael E. Raynor is a director in Deloitte Services LP. He leads the organization’s Center for Integrated Research. He is the coauthor, with Mumtaz Ahmed, of The Three Rules: How Exceptional Companies Think (New York: Penguin Books, 2013).


Derek Pankratz is a manager in the Center for Integrated Research in Deloitte Services LP.

Rob Del Vicario is a former manager in Deloitte Services LP.
Endnotes

1. Kelly E. See, Chip Heath, and Craig R. Fox, "Motivating individual performance with challenging goals: Is it better to stretch a little or a lot?" (working paper, New York University, 2003).
8. Ibid.
12. To "benchmark versus the market," we simply took the raw ROA and growth of US-based public companies active in 2013 and gave them a percentile rank from 1 to 100. To "benchmark versus peers," we calculated the simple five-year average of companies’ performance. We then broke them into peer groups based on SIC codes and size categories. Companies were assigned a percentile rank based on their place within their "bucket." For example, if a group has five companies, the top performer gets 100, the next highest gets 80, and so on. We used three-digit SIC codes to define industry-size groupings, unless the result was only a single company in the category, in which case we used two-digit SIC codes. The Compustat database is provided by Standard & Poor's.
13. The correlations between the "versus the market" and "versus peers" ranks and our method for growth are 0.29 and 0.43, respectively. The same correlations for ROA are 0.65 and 0.42.
14. The survey was in the field February 18–23, 2015. To participate, respondents needed to be in a director-level position or above at a US-based company with at least $500 million in revenue. Respondents also needed to indicate subject matter expertise in one or more areas relevant to the survey (e.g., finance, corporate strategy).