

Three years ago, Deloitte Consulting LLP launched The Persistence Project to identify the management practices that contribute most to sustained, superior corporate performance. Preliminary results have been published in the Harvard Business Review and the Annals of Applied Statistics. This article is the third in a series, providing a preview of the project's findings. See www.deloitte.com/us/persistence for more and to join the conversation.

RANK IGNORANCE

To get beyond the myths told of champions born of randomness we need an entirely new way to think about corporate performance

BY MICHAEL RAYNOR, MUMTAZ AHMED AND JEFF SCHULZ
> ILLUSTRATION BY SHOUT

Earnings seasons will soon be upon us, and we will be awash in explanations for the alleged feats and foibles of companies blessed or blighted enough to find themselves at the tip or tail of the roll call.

This annual rite is but the most visible example of a well-established tradition in the business press of imposing meaning on company performance. Whether it is earnings increases or decreases, the admiration of the business community, levels of innovation, share price appreciation or any number of other measures of noteworthy achievement, we are bombarded with explanations of why this or that company has ended up in this or that position on this or that list.

If this were nothing more than the C-suite's version of celebrity gossip site *TMZ* there would be no real harm done. And indeed it is a not uncommon conceit for many to believe, as with advertising, that the messages that bombard us might well affect others, but that we are immune to any ill effects.

Self-flattery is a poor defense against a cultural phenomenon. Companies with seemingly great performance along any of a number of different dimensions are the raw material for the ad hoc explanations that by dint of repetition become the aphorisms that inform the intuition of too many managers. Whether we recognize it or not, rankings of the biggest, fastest growing, most profitable, most admired, most innovative, best managed (we could go on), compiled by *Fortune*, *Forbes* and *Bloomberg BusinessWeek* provide all of us with the cast of characters from which we construct—consciously or otherwise—a narrative of the business environment around us.

It is standard procedure for compilers of these lists to make claims about the defining and distinctive attributes of top-ranked firms – after all, if we're not going to say something about why a firm ended up at number one, why bother having the list in the first place?



As a result, our interpretations of the world around us invariably begin to take on whatever hues and shades are implied by the inferred causes of the allegedly superior performance of these firms. Their management processes become leading practices, the opinions of their leaders become principles to live by, and any shortcomings are chalked up to the inevitable trade-offs attendant to success.

The problem is that none of the rankings that we are familiar with take into account the challenges associated with thinking about performance generally in a way that has a hope of revealing anything useful about the causes of superior performance of any kind. Assessing the value of a diamond requires an understanding the four Cs (cut, color, clarity and carat weight), and as it turns out, a clear-eyed examination of performance requires that we take explicit and careful account of at least the following five Cs of any alleged corporate gem:

- The **context** in which a given company has had to operate;
- The role of **chance** in a company's performance over time;
- Whether we are making a fair **comparison** among companies with different lifetimes;

- Whether a company's **course** of performance is stable, rising or falling; and
- Whether a company's trajectory of performance has **changed** over time.

CONTEXT

Most generally, the intent of performance rankings is to identify which firms have demonstrated exemplary performance. This might seem obvious, but this principle has implications that are often overlooked. For if we want to understand the relative performance of individual firms, we must strip away the influence of factors that obscure the contributions of firm-level attributes.

An intuitive example of the importance of non-firm attributes in determining firm-level performance can be seen in structural differences in innovation levels. Thanks to the visibility and frequency of innovations such as the iPod, the iPhone and the iPad, it is very tempting to rate Apple as more innovative than a company which operates in an industry where breakthroughs of a sort even analogous to those era-defining successes are typically separated by decades. But such differences can hardly be attributed to superior underlying innovativeness when such a deep structural industry-level characteristic differentiates the two firms. We cannot say anything about the relative innovativeness of these two companies until we condition each company's performance on what is possible given their respective contexts.

Those who prepare such rankings, like those who read them, often readily accept this principle. Yet, year after year, no meaningful attempt is made to distill the level of innovation driven by factors that a firm, and hence its management, can influence from those that are attributable to its environment (e.g., industry or markets served) or are otherwise beyond management's control (e.g., age or size at a point in time). We are simply told that such factors matter ... and then companies such as Apple and Google go at the top of the list.

If we do not explicitly and meaningfully control for these factors in a rigorous and consistent way—something no popular ranking algorithm does as far as we can tell—it is not merely tempting but nearly inevitable that as we seek the causes of superior performance we will end up making general principles out of what are fundamentally idiosyncratic and highly context-specific practices.

CHANCE

Even if confounding variables such as industry effects can be stripped away, perhaps the most insidious characteristic of ranked lists is the almost unavoidable inference that the firms at the top have achieved something noteworthy. It is standard procedure for compilers of these lists to make claims about the defining

and distinctive attributes of top-ranked firms – after all, if we’re not going to say something about *why* a firm ended up at number one, why bother having the list in the first place? Companies are complicit in this practice, pouncing on such rankings, at least when the results provide a pretext for making self-aggrandizing claims: There’s not a public relations department in creation that won’t trumpet having made the top 10 “best companies to work for” or “most environmentally friendly.” Commentators, corporate apologists and mere onlookers – we are all either subconsciously taken in by, or consciously exploiting, our seeming compulsion to believe that there is something special about being ranked above some defined set of peers.

The problem is that *every* list has someone in the top slot, and so simply being at the top is no guarantee that you’ve done something special or even that you’re any different from anyone else on, or for that matter off, the list.

To illustrate this phenomenon, Rebecca Henderson, a professor at Harvard Business School, begins her course with a lesson in the role of luck in determining outcomes. In her words:

I begin ... by asking all the students in the room to stand up. I then ask each of them to toss a coin: if the toss comes up “tails” they are to sit down, but if it comes up “heads” they are to remain standing. Since there are around 70 students in the class, after six or seven rounds there is only one student left standing. With the appropriate theatrics, I approach the student and say “HOW DID YOU DO THAT??!! SEVEN HEADS IN A ROW!! Can I interview you in Fortune? Is it the T shirt? Is it the flick of the wrist? Can I write a case study about you? ...

What’s at work here is our propensity to confuse the long-run consequences of systemic variability with individual attributes such as skill. It’s easy for us to make this mistake, because in any system subject to variation in outcomes—which is every system there is—streaks of high and low performance are to be expected, not due to any unique attributes of a given individual, but thanks to the inherent variability of the system.

In other words, luck can make it difficult to identify the truly great. Bad luck can lay low the mighty, while mediocre firms that catch enough lucky breaks can end up with seemingly exceptional performance. And so simply landing at the top of a rank ordered list based on recent metrics is not evidence of much of anything. It takes sustained superior performance to rise above the white noise of the system.

COMPARISONS

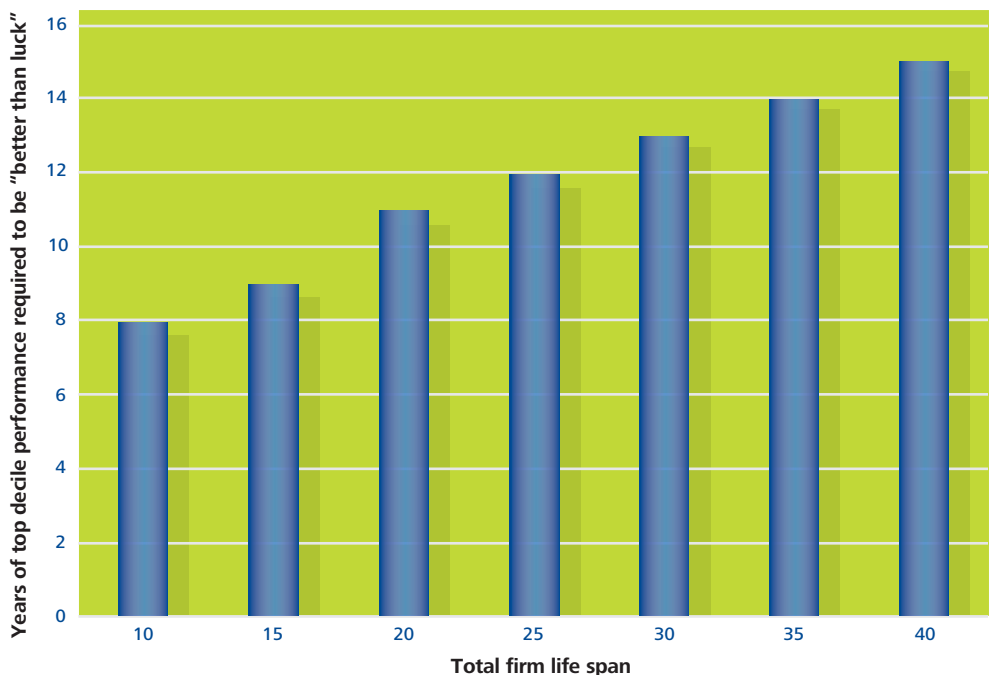
Similar to our comments on the importance of stripping out industry effects, the notion that short-run fluctuations can belie long-run trends is fairly widely

accepted. However, what is almost always overlooked is that what constitutes the long term can only be accurately assessed compared to any given firm's lifetime.

To see why, think of our coin tossing example. If I report that I've flipped five heads in a row, you might think that's a low probability event – just over a 3 percent chance, in fact. But if I then tell you that my five-in-a-row streak occurred as part of a sample of 10 flips, it's not nearly so remarkable, since I had six chances to flip five in a row, not just one. Similarly, taking, say, a five-year period of performance and picking off the winners does nothing to correct for short-run fluctuations and merely substitutes one selection bias for another.

To get a sense of the magnitude of this problem, consider two major research efforts that each see Campbell Soup as a company with noteworthy performance: *What Really Works* and *Big Winners and Big Losers*. The period under examination in the two studies overlaps by five years (from the beginning of 1992 to the end of 1996). Yet *What Really Works* (1986-1996) holds up Campbell Soup as a “winner”

Figure 1. The “long term” is relative



A firm with 10 years of data needs to deliver seven years in the top decile to have the same probability of being great as a firm with 40 years of data and 14 years in the top decile. Yet looking at randomly chosen five-year windows of exceptional performance would turn up our 40-year-old firm much more often. As a result, two firms with statistically indistinguishable track records would have very different probabilities of being identified as “great” – with potentially profound implications on our beliefs of what constitutes greatness.

while *Big Winners and Big Losers* (1992-2002) sees the company as a “big loser.” What’s going on?

Campbell Soup, across the sweep of its lifetime, has had its ups and downs. If you look at a “down” period, you’ll conclude the company is poorly managed; if you look at an “up” period, you’ll conclude it is well managed. And if by chance you happen to begin your analysis at an inflection point, you’ll conclude management is either brilliant or benighted.

This error is known generally as the “Texas Sharpshooter” problem, in which the target is defined only after the shots have been fired. In this context, it takes the form of focusing on specific periods of time that capture only a portion of most focal firms’ existences and separating out companies with suggestive performance profiles. Consequently, there is every reason to think that many or even most allegedly great results trumpeted in most list-making efforts owe more to Davey Crockett than to Peter Drucker.

COURSE

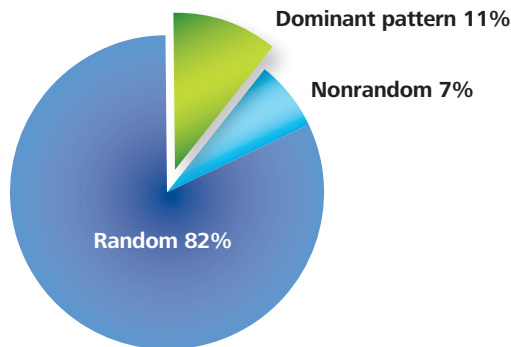
But wait! Isn’t it possible or even likely that characterizing a company’s performance over its entire life assumes a level of consistency that is vanishingly rare? Won’t most companies exhibit some sort of trend in their performance (either low to high or high to low), and so why can’t we learn something important about excellence by studying what they did when performing well?

Identifying firms with an exceptional level of performance over their lifetime means accounting for the variance of the system within which the firm operates. This is challenging enough. Determining whether or not a firm’s performance has a trajectory—flat, rising, falling, etc.—requires seeing a still finer-grained structure through the same blizzard of randomness. Since the only way to get more information out of any system is by analyzing more data, we require even longer periods of time and still stronger signals to reach valid conclusions.

Based on our analysis of U.S. public companies’ performance *trajectories* (rather than merely levels) we have concluded that barely 10 percent of the approximately 4,000 companies with the necessary minimum of 10 years of data have performance patterns characterized by a dominant trajectory. (Seven percent of firms have nonrandom profiles that stand out from the noise but not in ways that can be characterized by an identifiable trend.)

When we analyze the companies recently tagged by prominent publications as leading performers using this method, it turns out that only four of *Bloomberg BusinessWeek*’s top 40 global companies for 2009 exhibit identifiable patterns of performance, and of those four, two are deteriorating. Similarly, *Fortune*’s list of the

Figure 2. Firms with identifiable trajectories of performance are rare



50 most admired companies captures only 14 with meaningful patterns, of which six are in decline.

CHANGE

Finally, perhaps the most rhetorically compelling narrative in all of business is the “turnaround.” After the near-death experience triggers a corporate adrenaline rush, a new energy reinvigorates the entire organization. Paradigms are shifted, boxes are thought outside of, and there’s reinvention all around.

But if we thought it was difficult to see through the random noise to identify a caliber or a course of change, determining with any meaningful degree of certainty whether or not there are any changes within a single firm’s lifetime is more demanding still. With only one data point per year to work with, even the most sensitive statistical test and the most lenient definitions of “significant” typically require decades of data to reach a compelling conclusion.

Our analysis leads us to believe that, at some point in their histories, about 9 percent of the companies that have enough data to ground a credible opinion found their way onto a higher trajectory and another 9 percent fell to a lower trajectory. The vast majority of companies that dominate the high-profile lists we examined turn out to have, as with other dimensions of performance, a pattern that is indistinguishable from random walks.

KUDOS

The courts take great care to distinguish between “direct” and “indirect,” or circumstantial, evidence. Should you watch someone cut down your neighbor’s tree, you have direct evidence that they cut down the tree. Should you look out the window and see someone holding a chainsaw standing next to your neighbor’s felled tree, you have circumstantial evidence that they cut it down.

Publisher	List (2009)	No. of companies on list	No. of companies with nonrandom performance relevant to the list
<i>Bloomberg BusinessWeek</i>	World's Best Companies	40	7
<i>Fortune</i>	World's Most Admired Companies	50	12
<i>Barron's</i>	World's Most Respected Companies	100	7

For each of the lists identified above we count as “relevant” only those companies with non-random performance that is consistently strong, upward trending, or changing to an upward trending trajectory as relevant. Firms with statistically significant deteriorating performance are not counted since each of these lists seeks to capture firms with performance that is worthy of emulation.

It generally takes a great deal more circumstantial evidence than direct evidence to justify a conviction.

When it comes to thinking about what constitutes greatness in a company, we have not tended to be nearly so discriminating. The overwhelming tendency is to assume that performance “speaks for itself,” and companies with great performance are great companies simply because they have great performance. Our qualifications on this belief are limited only to some ill-defined sense of the “long term” and having obeyed the laws of the land.

This results in an unfortunate tendency to admire firms almost entirely on the basis of their economic performance and to seek to emulate the behaviors and traits of companies turning in the best numbers.

Our analysis suggests to us that raw performance is merely circumstantial evidence of organizational greatness. The world of corporate performance is a noisy system, and what might appear to be noteworthy, and perhaps even praiseworthy, is very often simply the consequence of luck.

This is not necessarily a problem when it comes to allocating economic benefits. Companies, and people, that succeed through luck have every right to the fruits of their good fortune. Since luck is, by definition, unpredictable and cannot be manipulated, its role in determining who wins and who loses, while it might offend our sense of justice, does nothing to undermine the fairness of competition. Since no one knows who will get lucky and who won't, we all still play the game the best we can. In fact, since luck is such a large factor in determining outcomes, it might even encourage all of us to try even harder: Even the mediocre have a real shot at the big time if we catch a few breaks, and almost no deficit is too much to

overcome if only a bounce or two goes your way.

Although we need not “correct” for luck when allocating economic reward it is worth considering the role of luck when allocating other valuable commodities, such as our moral approbation and our efforts to learn from the experiences of others. Those who win lotteries justifiably get to keep the money, but few of us admire lottery winners just because they won the lottery. (Don’t confuse envy with admiration!) Neither would any of us seek to learn anything about lottery winning from them. They are individuals operating within a system, and that system generates “winners” entirely independently of the individuals within it. Winning the lottery is its own reward.

In the corporate world, the rewards of great performance are similarly self-contained: Superior profitability or stock price appreciation is valuable. But companies with great performance do not merit our admiration or study merely by virtue of their great performance. For although competition among companies is not nearly so dominated by luck as lotteries, neither is it a perfect meritocracy.

In short, until we approach the analysis of corporate performance with much greater subtlety we will be unable to see through the “Hollywood” associated with rankings of various sorts. We must school our intuitions to accept the fact that the world of corporate performance is a noisy system. Changes in this year’s earnings may be nothing more than a regression to the mean or stochastic beneficial shocks, and so indicative only of properties of the larger system, not any attributes of an individual firm.

(This is something many people say they believe, but how many of you will look at your own organization’s performance this year and offer up an explanation that points either to the malevolence of your circumstances or your organization’s inherent worth depending on the outcome?)

Beyond simply being honest with ourselves about the significance of short-term fluctuations, only by insisting on drawing conclusions about the drivers of great performance by studying those firms that are demonstrably great can we make true progress. And once we are able to identify the behaviors associated with greatness, it will be possible not only to identify it sooner, but also perhaps to deliberately create it.

Michael Raynor is a director in the Strategy & Operations practice of Deloitte Consulting LLP. He is the co-author of The Innovator’s Solution and the author of The Strategy Paradox.

Mumtaz Ahmed is a principal in the Strategy & Operations practice of Deloitte Consulting LLP.

Jeff Schulz is a consultant in the Strategy & Operations practice of Deloitte Consulting LLP.

The authors can be reached at persistenceproject@deloitte.com.