Great Companies: *Looking for Success Secrets in All the Wrong Places*

GABRIELLE BAUM AND GARY SMITH

GABRIELLE BAUM
is an investment banking analyst at Deutsche Bank in New York, NY.
gabbaum422@gmail.com

GARY SMITH
is a professor in the Department of Economics at Pomona College in Claremont, CA.
gsmith@pomona.edu

Why are some companies more successful than others? Why do some companies grow and prosper while others languish and fail? Why are some companies great while others are merely good, mediocre, or bad? These questions are repeatedly asked by business executives, management consultants, and financial analysts, but the answers remain elusive.

In this article, we contrast the companies identified as successful by the best-selling books *Good to Great* (Collins [2001]) and *In Search of Excellence* (Peters and Waterman [1982]) with *Fortune* magazine's most-admired surveys. First, we discuss their selection criteria. Then we look at how well these companies' stocks did after the publication of these lists.

The history of corporate America has witnessed the rise and fall of many companies and even entire industries. In the early 1990s, having a dot-com at the end of a company's name was assumed to guarantee success, and stock prices soared to levels that turned out to be completely unjustified. A dot-com looked like the formula for success, but it was not.

Stock prices are often pumped up by dreams of reaping riches from being part of the latest fad or the next new thing—the next IBM, Microsoft, or Google. Are there any objective criteria for predicting which companies will be successful and which stocks will be profitable investments? There are three major pitfalls facing those who would predict success: data mining, the efficient market hypothesis, and regression toward the mean.

DATA MINING

Statistical tests assume that a researcher starts with a theory, collects data to test the theory, and reports the results—whether statistically significant or not. However, many people work in the other direction, scrutinizing the data until they find a pattern and then formulating a theory that fits the pattern. Ransacking data for patterns is fun and exciting—like playing Sudoku or solving a murder mystery. Examine the data from every angle; look for something, anything that is interesting; after a pattern is discovered, think about reasons behind it.

This pillaging is known as *data mining* (or data grubbing, data dredging, fishing expedition). The problem with this approach is that even random coin flips form patterns that appear to be meaningful but are, in fact, meaningless. When a fair coin is flipped 10 times, a streak of four heads in a row (or four tails in a row) seems too remarkable to be explained by chance, although streaks this long, or longer, can be expected 47% of the time. One may think something must be unusual about the coin or the person flipping

the coin and thus underestimate the importance of luck. This is instead another version of the Texas sharpshooter fallacy, in which a person with no talent for shooting fires randomly at the side of a barn and, afterward, paints a bullseye around the cluster of bullet holes.

Even randomly generated data typically contain clusters and, should an explanation for a particular cluster be sought, it will inevitably be found. In a cancer study, for example, one might discover that several cancer victims happened to live near power lines, a Little League field, or a water tower—which proves nothing at all. Data mining demonstrates little more than a researcher's endurance. Data without theory is treacherous, and one should be deeply skeptical of any supposed results gathered through data mining.

The Foolish Four investment strategy was introduced in Gardner and Gardner's [1996] book *The Motley Fool Investment Guide: How the Fools Beat Wall Street's Wise Men and How You Can Too*. The Foolish Four Strategy operates as follows:

1. Identify the 10 Dow Jones Industrial stocks with the highest dividend yields (dividend/price) at the beginning of each year.
2. Of these 10, identify the five stocks with the lowest prices.
3. Drop the stock with the lowest price.
4. Invest 40% of your wealth in the next-to-lowest priced stock and 20% each in the other three stocks.

The Gardners reported that during the years 1973–1993, the Foolish Four strategy had an annual average return of 25% and that it "should grant its fans the same 25 percent annualized returns going forward that it has served up in the past" (p. 104).

There is a glimmer of logic here in that a contrarian strategy may select stocks with low prices relative to dividends (or earnings or assets). Beyond that, it is pure data mining. There is no reason why the per share price should matter because price per share depends on the number of shares outstanding; why the lowest-priced stock should be dropped; or why a double weight should be given to the second-lowest-priced stock.

McQueen and Thorley [1999] found the Foolish Four strategy unimpressive both during the presample period (1949–1972) and during the in-sample period if the strategy was implemented on the first trading day of July rather than the first trading day of January. If the strategy had any true merit, it would not be as sensitive to the choice of time period or starting month.

In 1997, one year after the introduction of the Foolish Four, the Gardners adjusted the system and renamed it the UV4. Their explanation for doing so confirms the usage of data mining: "Why the switch? History shows that the UV4 has actually done better than the old Foolish Four" (Sheard [1997]). It is hardly surprising that a data-mined strategy fails to perform as well out of a chosen sample as it did in the chosen sample. The Gardners admitted as much when the Motley Fool stopped recommending the Foolish Four and UV4 strategies in 2000 (Gardner and Gardner [2000]).

THE EFFICIENT MARKETS HYPOTHESIS

In an efficient market, all available information is taken into account by investors and is therefore fully reflected in market prices (Fama [1991]). If it is well known that a company is exceptionally strong, then its stock will trade at a price that gives investors an appropriate anticipated return, taking into account risk and other characteristics that are relevant for investment decisions.

There is, however, a difference between possessing information and processing information. Warren Buffett did not excel in the market for decades by having access to information that was not available to other investment professionals, but rather by thinking more clearly about information available to everyone. The question this article poses is whether best-selling books like *Good to Great* and *In Search of Excellence* examine companies differently than others and therefore uncover strengths that are not appreciated fully by security analysts and other investors.

Do successful companies equate to great investments? Collins explicitly identifies great companies by their exceptional stock market performance; if their greatness is real and not simply an illusion resulting from data mining of serendipitous stock market fluctuations, then these companies' exceptional stock market performance should continue. Does it do so?

REGRESSION TO THE MEAN

Horace Secrist had a distinguished career as a professor of economics at Northwestern University. He

wrote 13 textbooks and was director of Northwestern's Bureau of Economic Research. In 1933, in the depths of the national economic tragedy that became known as the Great Depression, he published a book titled *The Triumph of Mediocrity in Business* [1933] that he hoped would explain the causes, provide solutions, and secure his legacy.

In every industry that he examined, the most successful companies tended to become less successful over time, whereas the least successful tended to become more successful. Secrist reached the conclusion that all companies would soon be mediocre. His explanation was as follows:

> Complete freedom to enter trade and the continuance of competition mean the perpetuation of mediocrity. New firms are recruited from the relatively "unfit"—at least from the inexperienced. If some succeed, they must meet the competitive practices of the class, the market, to which they belong. Superior judgment, merchandizing sense, and honesty, however, are always at the mercy of the unscrupulous, the unwise, the misinformed, and the injudicious. The results are that retail trade is overcrowded, shops are small and inefficient, volume of business inadequate, expenses relatively high, and profits small. So long as the field of activity is freely entered, and it is; and so long as competition is "free," and, within the limits suggested above, it is; neither superiority nor inferiority will tend to persist. Rather mediocrity tends to become the rule (p. 24).

The nation's economic problems were evidently a result of the new economic principle Secrist had discovered: Competitive pressures inevitably dilute superior talent. The apparent solution to this problem was to protect superior companies from the competition of less-fit companies attempting to enter the market.

The president of the American Statistical Association wrote an enthusiastic review of Secrist's work, as did the *Journal of Political Economy, Journal of the Royal Statistical Society, American Economic Review, and Annals of the American Academy of Political and Social Science*. These reviewers and Secrist alike were all deceived by regression toward the mean.

To understand regression, one should imagine that 100 people are asked 20 questions about the stock market. Each person's "ability" is then his or her average score on an infinite number of such tests. Some people have an ability of 90, some 80, and some near zero. A person with an ability of, for example, 80 will not score 80 on *every* test. Therefore, what can be inferred from a person's score on a single test? A person who scores in the 90th percentile on a test could be someone of more modest ability who performed unusually well; this person could also be someone of higher ability who did poorly on this particular test. The former is more likely because there are more people who are below the 90th percentile than are above it. If this specific person's ability is, in fact, below the 90th percentile, then his or her score on subsequent tests will probably likewise be below the 90th percentile. Similarly, a person who scores far below average is likely to have had a less successful day than usual and will probably score somewhat higher on later tests. This tendency of people who obtain results that are far from the mean to fall closer to the mean on second testing is an example of regression toward the mean.

Regression occurs in many contexts. School children who are given special tutoring to better their low test scores can be expected to do better on subsequent tests even if the tutors do nothing more than snap their fingers. Patients who are given treatment because of a worrisome result on a medical test can be expected to improve even if the treatment is worthless. An athlete who wins a rookie-of-the-year award can be expected to have a sophomore slump.

In the 1800s, Sir Francis Galton observed that unusually tall parents tend to have somewhat shorter children, with the reverse being true of unusually short parents. The conclusion that heights are regressing to mediocrity—a conclusion encouraged by Galton himself in titling his study "Regression Towards Mediocrity in Hereditary Stature"—is erroneous. Regression toward the mean does not imply that every person will eventually be the same height any more than it implies that everyone will soon achieve the same score on tests. What regression does imply, however, is that imperfect measurement of an unobserved trait tends to overstate the extent to which the underlying trait varies from the mean.

This framework is directly applicable to any imperfect measure of a company's success. For example, a company experiencing earnings growth that is high relative to a group of companies is also likely to be high relative

to that particular company's "ability." We can consequently anticipate the regression toward the mean of the company's earnings growth. Fama and French found such an earnings regression, although they attributed it to competitive forces rather than the purely statistical explanation that companies with relatively high earnings are likely to have experienced more good luck than bad:

> In a competitive environment, profitability is mean reverting within as well as across industries. Other firms eventually mimic innovative products and technologies that produce above normal profitability for a firm. And the prospect of failure or takeover gives firms with low profitability incentives to allocate assets to more productive uses (Fama and French [2000], p. 161).

This quotation is reminiscent of a book (Baumol et al. [1989]) and book review (Williamson [1991]), both written by eminent economists, which argued that nations' economic growth rates converge over time. The authors of both pieces completely ignored the role of regression to the mean in this convergence, an error that Milton Friedman [1992] discussed in a commentary aptly titled "Do Old Fallacies Ever Die?"

If finance professors and prominent economists can ignore regression to the mean, so, too, can anyone. There is well-established evidence that regression to the mean is a pervasive but subtle statistical principle that is often misunderstood or insufficiently appreciated (Kahneman and Tversky [1973]). Regarding the stock market, Keynes [1936] observed that "day-to-day fluctuations in the profits of existing investments, which are obviously of an ephemeral and nonsignificant character, tend to have an altogether excessive, and even absurd, influence on the market" (p. 340).

A large body of empirical literature on the topic of mean reversion in stock prices and stock returns exists. Some of the suggested explanations—noise trading (Poterba and Summers [1988]), fads (Shiller [1984]; Summers [1986]; McQueen [1992]), and overreaction to financial news (De Bondt and Thaler [1985, 1987])—are consistent with regression to the mean. An explanation based on rational investor reactions to competitive forces (Fama and French [1988, 2000]) is not consistent with regression to the mean, and some studies lack any theoretical underpinning at all (Kim et al. [1991]; Richardson [1993]).

Lakonishok et al. [1994] found that companies that have been performing well, as gauged by such metrics as earnings growth rates, in actuality are relatively poor investments compared with companies doing poorly by the same metrics. These results are consistent with the idea that investors do not fully appreciate regression and consequently are surprised when earnings do not grow as fast as in they have in the past. This surprise, in turn, has a negative effect on the company's stock price.

La Porta [1996] found that stocks of companies predicted to have high rates of earnings growth in reality underperform compared with stocks of companies about which analysts are more pessimistic. Keil et al. [2004] similarly found that earnings forecasts systematically fall to extremes, being too optimistic regarding companies predicted to do well and too pessimistic about those predicted to do poorly. The accuracy of these forecasts can be improved consistently and substantially by reducing their variance from mean forecast. Keil et al. also found that a portfolio containing stocks from pessimistic companies systematically outperformed an optimistic portfolio.

The same is true of virtually any measure of a company's success. The most successful companies have likely benefitted from luck and will consequently regress to the mean. If investors do not anticipate this regression, stock prices that are initially too high will adjust after the regression occurs.

GOOD TO GREAT

In 2001, Jim Collins published his best-selling book on management, titled *Good to Great: Why Some Companies Make the Leap... And Others Don't*. It has sold more than 4 million copies and appeared on several lists as among the best management books of all time. Collins wrote that his book reflects "our search for timeless, universal answers that can be applied by any organization." His conclusion is that "We believe that almost any organization can substantially improve its stature and performance, perhaps even become great, if it conscientiously applies the framework of ideas we've uncovered" (Collins [2001], p. 5).

Collins and his research team spent five years researching the 40-year stock market history of 1,435 companies and identified 11 stocks that outperformed the overall market and continued to improve 15 years after making the leap from good to great:

- Abbott Laboratories
- Circuit City
- Fannie Mae
- Gillette
- Kimberly-Clark
- Kroger
- Nucor
- Philip Morris
- Pitney Bowes
- Walgreens
- Wells Fargo

Collins scrutinized these 11 great companies and identified five common themes that he labeled with catchy names:

1. Level 5 Leadership: Having leaders who are personally humble, but professionally driven to make a company great.
2. First Who, Then What: Hiring the right people is more important than having a good business plan.
3. Confront the Brutal Facts: Good decisions take into account all of the facts.
4. Hedgehog Concept: It is better to be a master of one trade than a jack of all trades.
5. Build Your Company's vision: Adapt operating practices and strategies, but do not abandon the company's core values.

These characteristics are plausible and the names are memorable. The problem with Collins' work is that it is a backward-looking study that is undermined by data mining. Collins wrote that "we developed all of the concepts in this book by making empirical deductions directly from the data. We did not begin this project with a theory to test or prove. We sought to build a theory from the ground up, derived directly from the evidence "(Collins [2001], p. 10)".

Collins seemed to think this statement made his study appear unbiased and professional. He did not simply craft his conclusions: He went wherever the data led. In reality, Collins was admitting that he had no idea why some companies do better than others, and he was revealing that he was blissfully unaware of the perils of deriving theories from data. When considering any group of companies at a particular moment in time, whether the companies are the best or the worst of the

group, commonalities can always be found. For example, every company among the 11 selected by Collins has either a letter *i* or *r* in its name, and several have both an *i* and an *r*. Is ensuring that the company's name has an *i* or *r* in it the key to improving from a good company to a great company? Of course not.

Finding a pattern like the preceding is an obvious example of data mining. So, too, is examining a sequence of coin flips and noting that two heads happened to be followed by a tail more than 50% of the time. Collins' data mining is less obvious because his unearthed patterns sound like a plausible theory, but it is data mining nonetheless because, as Collins freely admits, he created his theory after looking at the data.

To buttress the statistical legitimacy of his theory, Collins spoke to two professors at the University of Colorado. One said that "the probabilities that the concepts in your framework appear by random chance are essentially zero." The other professor was more specific. He was asked, "What is the probability of finding by chance a group of 11 companies, all of whose members display the primary traits you discovered while the direct comparisons do not possess those traits?" The professor calculated this probability to be less than 1 in 17 million. Collins concludes, "There is virtually no chance that we simply found 11 random events that just happened to show the good-to-great pattern we were looking for. We can conclude with confidence that the traits we found are strongly associated with transformations from good to great" (Collins [2001], p. 212).

It is not clear how this probability of 1 in 17 million was calculated. (I contacted the professor and he could not remember.) What is clear is that the calculation is incorrect. In statistics, this kind of reasoning is sometimes called the Feynman Trap, a reference to the Nobel Laureate Richard Feynman. Feynman asked his Cal Tech students to calculate the probability of his leaving the classroom and finding that the first car in the parking lot had a specific license plate; for instance, 8NSR26. Cal Tech students are very smart and they quickly calculated a probability by assuming each number and letter were independently determined: less than 1 in 17 million. When they finished, Feynman revealed that the correct probability was 1 because he had seen this license plate on his way to class. Something extremely unlikely is not unlikely at all if it has already happened. The calculations made by the Colorado professors and the Cal Tech students assume that the five traits and the license plate

number were specified before looking at the companies and the cars. They were not, and these calculations are therefore irrelevant.

Collins does not provide any evidence that the five characteristics he describes were responsible for these companies' success. To do so, he would need to provide a theoretical justification for these characteristics, select companies before beginning his study that did and did not have these characteristics, and monitor their success according to some metric also established previously. He did none of this.

After the publication of *Good to Great*, success evaporated for some of these companies. Fannie Mae stock went from above $80 a share in 2001 to less than $1 a share in 2008 and subsequently delisting in 2010. Circuit City went bankrupt in 2009. Niendorf and Beck [2008] and Resnick and Smunt [2008] both concluded that, overall, the out-of-sample stock returns for these 11 companies in the years 2005 and 2006 did not show substantial or statistically significant abnormal returns.

IN SEARCH OF EXCELLENCE

Twenty years before Collins' book, another best-selling book on business did something very similar and experienced exactly the same problems. Two McKinsey consultants, Tom Peters and Robert Waterman, were asked to study several successful companies. They spoke with other McKinsey consultants and devised a list of 62 leading companies.

To make their analysis appear scientific, they looked at six measures of long-term success, three related to growth and three measuring the return on capital and assets. To maintain its place in the sample, a company had to rank in the top 50% of its industry for four of the six measures during the period 1961–1980. As a final screening measure, Peters and Waterman asked industry experts to rate the companies' 20-year record of innovation. The final 43 firms included the 36 publicly traded companies shown in Exhibit 1 and eight companies that were privately held or subsidiaries of other companies: Allen-Bradley, Atari, Bechtel, Chesebrough-Pond's, Frito-Lay, Hughes Aircraft, Mars, and Tupperware.

Peters and Waterman then spoke to managers and read magazine stories about these companies. They uncovered eight common traits; for example, a bias for action and being close to the consumer. The book they wrote

EXHIBIT 1
The 36 Publicly Traded Companies Identified in *In Search of Excellence*

3M	Disney Productions	Marriott
Amdahl	Dow Chemical	Maytag
Amoco	Du Pont	McDonald's
Avon	Eastman Kodak	Merck
Boeing	Emerson Electric	National Semiconductor
Bristol-Myers	Fluor	Proctor & Gamble
Caterpillar Tractor	Hewlett-Packard	Raychem
Chesebrough-Pond's	IBM	Revlon
Dana Corporation	Intel	Schlumberger
Data General	Johnson & Johnson	Texas Instruments
Delta Airlines	K Mart	Walmart
Digital Equipment	Levi Strauss	Wang Labs

about their efforts, *In Search of Excellence* [1982], was again a backward-looking study undermined by data mining. There is no way of knowing whether companies with "a bias for action" were more successful than other companies or if past excellence indicated future excellence.

Clayman [1987, 1994] conducted two studies of the stock returns of companies labeled excellent by Peters and Waterman. In the first study, Clayman looked at 29 of the 36 publicly traded companies over the period 1981–1985. Eighteen underperformed the stock market, but some of the 11 companies that outperformed the market (including Walmart, Maytag, and McDonald's) did very well, so that an equally weighted portfolio exceeded the market by 1.1% per year. In the second study, Clayman [1994] concluded that, for the years 1988–1992, the investment performance of the 29 excellent companies as a whole was indistinguishable from that of "unexcellent" companies.

Clayman's explanation is an incorrect interpretation of regression to the mean, similar to that of Fama and French [2000]:

> Over time, company results have a tendency to regress to the mean as underlying economic forces attract new entrants to attractive markets and encourage participants to leave low-return businesses. Because of this tendency, companies that have been "good" performers in the past may prove to be inferior investments, while "poor" companies frequently provide superior investment returns in the future. (Clayman [1987], p. 63)

Regression to the mean does not assume that "economic forces tend to move things towards equilibrium" (Clayman [1987], p. 62). Economic forces may exist, but regression to the mean is a purely statistical phenomenon that occurs when observed data measure unobserved traits imperfectly; for example, using observed earnings growth rates to measure a company's greatness.

Bannister [1990] found that the stock returns for non-excellent companies (i.e., companies in the bottom third of the six screening categories used by Peters and Waterman) generally exceeded the stock returns for excellent companies (those in the top third in all six categories) during the years 1977–1989. His explanation echoes Clayman's incorrect interpretation of regression to the mean: "High returns eventually invite new entrants, driving down profitability, while poor returns cause the exit of competitors, leaving a more profitable industry for the survivors" (Bannister [1990], p. 68).

FORTUNE'S MOST ADMIRED LIST

Since 1983, *Fortune* has published an annual list of "America's Most Admired Companies." The differences between this list and the lists appearing in many management books is instructive. The *Fortune* list is based on a survey of thousands of business executives, directors, and analysts who rate the largest companies in their industry on eight key attributes that *Fortune*'s editors believe are crucial for success: innovation, people management, use of assets, social responsibility, management quality, financial soundness, long-term investment, and product quality. In recent years, a ninth attribute (global competitiveness) has been added to the list. One virtue of the *Fortune* list is that it is based on this set of ex ante criteria. The same measures of successful companies are used year after year. Instead of looking for ex post traits that are common to admired companies, *Fortune* seeks companies that have predefined traits.

Although the list is compiled annually, the people being surveyed evidently consider the companies' long-term characteristics, not the most recent trend in their fortunes. On average, 72% of the companies in the top 10 one year are in the top 10 the following year. In the thirty years from 1983 to 2013, Procter & Gamble has been in the top 10 a total of 19 times and Walmart 14 times. Berkshire-Hathaway made the top 10 for 17 consecutive years, from 1997 to 2013.

Antunovich et al. [2000] found that, over the period 1983–1995, the stocks of the companies in the top decile of the *Fortune* ratings performed better than the stocks in the bottom decile. Using data from 1983 to 2004, Anderson and Smith [2006] found that a portfolio consisting of the stocks of the 10 most-admired companies outperformed the market, whether the stocks were purchased on the publication date or 5, 10, 15, or 20 trading days later. This is a clear challenge to the efficient market hypothesis considering the *Fortune* list is readily available public information. Anderson and Smith [2006] conclude that "We have no compelling explanation for this anomaly. Perhaps Philip Fisher was right: the way to beat the market is to focus on scuttlebutt—those intangibles that don't show up in a company's balance sheets—and *Fortune*'s most-admired survey is the ultimate scuttlebutt" (p. 93).

Using data for 1983 through 2007, Anginer and Statman [2010] confirmed that a portfolio consisting of the 10 companies with the highest *Fortune* ratings did well relative to other admired stocks and relative to the 10 least admired stocks, but they also found that, overall, the stock performance was better for the companies in the bottom half of the ratings than for companies in the top half, consistent with regression to the mean. The superior performance of the top 10 companies remains an anomaly.

METHODS

Earlier studies that examined the out-of-sample performance of *Good to Great* stocks (Niendorf and Beck [2008]; Resnick and Smunt [2008]) and *In Search of Excellence* stocks (Clayman [1987, 1994]; Bannister [1990]) often used time periods as brief as four years and either ignored risk or gauged risk by the capital asset pricing model (CAPM). In contrast, it is now considered almost mandatory to use the Fama–French factor model to test for abnormal returns.

We looked at the daily returns of the 11 *Good to Great* stocks (Great) and the 35 publicly traded *In Search of Excellence* stocks (Excellent) from the first trading day in January after the publication of each book through December 31, 2012. We also looked at a portfolio consisting of an equal investment in each of the 11 Great stocks and a similar portfolio for the 35 Excellent stocks. If a company was acquired or became private, the proceeds were evenly invested in the remaining stocks in

the portfolio. For example, after Gillette merged with Proctor & Gamble on October 1, 2005, the proceeds of the merger were invested equally in the 10 stocks remaining in the Great portfolio.

We looked at the risk-adjusted returns for the individual stocks and for each portfolio using the Fama–French three-factor model, augmented by a momentum factor:

$$R = \alpha + \beta_1 MKT + \beta_2 SMB + \beta_3 HML + \beta_4 UMD + \varepsilon$$

in which R = return on an individual stock or portfolio minus return on Treasury bills; MKT = NYSE/AMEX/NASDAQ market index return minus return on Treasury bills; SMB = average return on three small portfolios minus the average return on three big portfolios (size factor); HML = the average return on two value portfolios minus the average return on two growth portfolios (book-to-market factor); UMD = average return on two high prior return portfolios minus the average return on two low prior return portfolios (momentum factor). The data were downloaded from Center for Research in Security Prices and Ken French's website (French [2013]).

The Fama–French three-factor model accounts for the fact that, in general, small stocks tend to outperform big stocks (Banz [1981]; Reinganum [1981]) and value stocks with high book-to-market ratios tend to outperform growth stocks (Rosenberg et al. [1985]). A momentum factor is included because of empirical evidence that stocks that have been doing well tend to outperform those doing poorly (Jegadeesh and Titman [1993]).

The use of three portfolios for SMB and two portfolios for HML (and UMD) suggests data mining, but nonetheless, the Fama–French model has become standard for gauging abnormal returns. It is undecided whether these factors reflect risks that matter to investors (Chan [1988]; Fama and French [1992]) or are evidence of market inefficiencies (Lakonishok et al. [1994]). Nonetheless, the question under consideration is whether the performance of the Great and Excellent stocks can be explained by these four factors; a substantial, statistically significant value for alpha indicates an unexplained excess return.

A similar analysis was done for Fortune's Most Admired stocks [2015] (Admired). Following Anderson and Smith [2006], the Admired strategy invests equally

in each of the 10 most admired stocks on the printed publication date in 1983. Each year thereafter, the portfolio is liquidated on that year's publication date and the proceeds are invested in the current year's most admired companies. Investors can easily implement this strategy because the magazine appears roughly one week before the publication date printed on its cover.

Because Fortune's Most Admired list is compiled annually, the Admired portfolio does not gauge the longer-term performance of the Most Admired stocks. To investigate this topic, we averaged the returns across the annual portfolios, beginning on each year's publication date. Thus, we looked at the daily returns for each year's top-10 companies and the market index, starting on that year's publication date and continuing until December 31, 2012. We then calculated the average returns for the Most Admired Stocks and the market index on the first trading day after the publication date, then second trading day, and so on until the last trading day in the dataset.

RESULTS

Exhibit 2 shows the annual returns for each Great stock as compared with the index, the returns of which are not the same for each stock because of differences in holding periods due to mergers, acquisitions, buyouts, and bankruptcies. Five of the Great stocks did better than the market; six did worse. For the Fama–French factor model, the only alphas with P-values less than 0.05 were for Philip Morris (a positive alpha) and Circuit City (a negative alpha).

EXHIBIT 2
Annual Returns and Fama–French Alphas, *Good to Great Stocks*, January 2, 2002–December 31, 2012

	Annual Returns (%)		Four-Factor Model	
	Stock	Index	Alpha	P-Value
Abbott Laboratories	4.88	5.23	0.017	0.465
Circuit City	–100.00	–0.03	–0.219	0.036
Fannie Mae	–100.00	2.80	–0.139	0.268
Gillette	17.89	6.03	0.070	0.088
Kimberly-Clark	6.76	5.23	0.017	0.334
Kroger	2.99	5.23	0.008	0.773
Nucor	14.62	5.23	0.049	0.197
Philip Morris	17.09	5.23	0.057	0.027
Pitney Bowes	–6.43	5.23	–0.037	0.111
Walgreens	2.08	5.23	0.004	0.886
Wells Fargo	7.04	5.23	0.020	0.549

Exhibit 3 shows the results for the Excellent stocks. Fifteen of the Excellent stocks performed better than the market, and 20 fared worse. The only alphas with P-values less than 0.05 were Walmart and Intel (both positive) and Delta and Dana (both negative).

Exhibit 4 shows the annual returns and alphas for the three portfolios. In contrast to the Great and Excellent portfolios, which slightly underperformed the market, the Admired portfolio had an annual return of 12.57% compared with 10.39% for the CRSP index. All three alphas are positive, but only the Admired portfolio has a statistically significant alpha. The success of the Admired portfolio does not appear attributable to

EXHIBIT 3

Annual Returns and Fama–French Alphas, *In Search of Excellence* Stocks, January 3, 1983–December 31, 2012

	Annual Returns (%)		Four-Factor Model	
	Stock	Index	Alpha	P-Value
3M	11.16	10.49	0.011	0.416
Amdahl	4.21	16.27	–0.000	0.993
Amoco	16.80	16.44	0.005	0.814
Avon	8.86	10.49	0.010	0.650
Boeing	11.70	10.49	0.016	0.373
Bristol-Myers	11.18	10.49	0.019	0.262
Caterpillar Tractor	12.37	10.49	0.015	0.436
Dana Corporation	–100.00	12.40	–0.091	0.007
Data General	0.67	16.13	0.028	0.589
Delta Airlines	–100.00	12.10	–0.111	0.003
Digital Equipment	0.84	16.38	–0.008	0.827
Disney Productions	13.88	10.49	0.025	0.164
Dow Chemical	10.13	10.49	0.003	0.858
Du Pont	10.83	10.49	0.007	0.655
Eastman Kodak	–100.00	10.47	–0.064	0.058
Emerson Electric	11.17	10.49	0.007	0.608
Fluor	8.58	10.49	–0.001	0.975
Hewlett-Packard	5.66	10.49	0.022	0.338
IBM	9.39	10.49	0.022	0.159
Intel	14.75	10.49	0.080	0.009
Johnson & Johnson	13.31	10.49	0.026	0.068
K Mart	–100.00	11.65	–0.081	0.066
Levi Strauss	15.69	15.72	–0.024	0.758
Marriott	11.05	14.64	0.005	0.894
Maytag	7.17	12.45	–0.007	0.792
McDonald's	14.62	10.49	0.029	0.075
Merck	13.36	10.49	0.027	0.109
National Semi	7.18	10.02	0.064	0.064
Proctor & Gamble	12.99	10.49	0.021	0.151
Raychem	8.54	16.17	0.000	0.992
Revlon	30.04	17.52	0.046	0.428
Schlumberger	8.41	10.49	–0.009	0.684
Texas Instruments	9.53	10.49	0.046	0.064
Walmart	17.17	10.49	0.066	0.002
Wang Labs	–100.00	14.55	–0.069	0.554

EXHIBIT 4

Three Portfolios

	Annual Returns (%)		Fama–French	
	Stock	Index	Alpha	P-Value
Good to Great	5.20	5.23	0.006	0.578
In Search of Excellence	10.29	10.49	0.010	0.121
Most Admired	12.57	10.39	0.019	0.001

the effects of the market, size, value, or momentum factors.

Exhibit 5 shows the estimated coefficients for the four-factor model with the t-values in brackets. For all three portfolios, the coefficients of the market factor are less than one and the coefficient of the SMB factor is negative, which is consistent with the relatively large size of these companies. The negative coefficient for the HML factor in the Excellent and Admired portfolios is consistent with Fama and French's [1995] conclusion that strong firms with consistently strong earnings tend to have negative HML coefficients. However, the HML coefficient for the Great portfolio is positive. The momentum factor is positive for the Great and Excellent portfolios, indicating that these stocks benefit from momentum, but the momentum coefficient is negative for the Admired portfolio.

Exhibit 6 compares the performance of the Most Admired stocks with the market index over long horizons. There are approximately 250 trading days in a calendar year, so the results in Exhibit 6 are presented in roughly two-year intervals stretching up to 30 years. For example, 20 years after being selected as Most Admired companies, wealth grew, on average, from 1 to 10.124 for the Admired companies and to 7.197 for the market index.

There are fewer portfolios in the longer horizons, with the limit being one portfolio (the 1983 list) with a 30-year horizon. The Most Admired stocks outperformed the market index over every horizon with P-values less than 0.05, except for the very longest horizons for which there is relatively little data.

Exhibit 7 shows the ratio of the Most Admired wealth to the index wealth. This ratio is more volatile for the longest horizons and includes the least data but nonetheless demonstrates that, on average, the Most Admired stocks outperformed the market over long horizons.

EXHIBIT 5
Estimated Four-Factor Model

	Alpha	MKT	SMB	HML	UMD	R²
Good to Great	0.0062	0.9274	−0.1403	0.1084	0.0607	0.815
	[0.56]	[98.08]	[7.22]	[5.10]	[4.76]	
Search of Excellence	0.0096	0.7022	−0.2241	−0.1128	0.0429	0.701
	[1.55]	[118.60]	[20.68]	[9.30]	[5.26]	
Most Admired	0.0191	0.9440	−0.2746	−0.3342	−0.0955	0.823
	[3.08]	[159.06]	[25.15]	[27.32]	[11.62]	

Note: t-value in square brackets.

EXHIBIT 6
Wealth Across Different Horizons for the Most Admired Companies

| Number of Trading Days	Years (approximate)	Number of Portfolios	Most Admired Wealth	Index Wealth	P-Value
500	2	28	1.334	1.251	0.045
1,000	4	26	1.756	1.519	0.013
1,500	6	24	2.156	1.878	0.048
2,000	8	22	2.701	2.352	0.038
2,500	10	20	3.540	2.955	0.009
3,000	12	18	4.876	3.766	0.005
3,500	14	16	6.450	4.816	0.005
4,000	16	14	8.506	5.973	0.005
4,500	18	12	9.697	6.640	0.003
5,000	20	11	10.124	7.197	0.003
5,500	22	9	12.688	8.993	0.002
6,000	24	7	15.806	10.660	0.009
6,500	26	5	17.147	11.361	0.047
7,000	28	3	22.774	15.071	0.147
7,500	30	1	20.809	19.683	N/A

N/A = not applicable.

CONCLUSION

The secrets for building a successful company are still a secret, which is unsurprising. If the characteristics enumerated in *Good to Great, In Search of Excellence* and similar books were a guaranteed recipe for success, every company would be great. The real lesson from the enduring popularity of such advice is that the authors who write these books and the millions of people who buy them do not realize that the books are fundamentally flawed. This problem plagues the entire genre of formulas, secrets, and recipes for a successful business, a lasting marriage, living to the age of 100, and so on that are based on backward-looking studies of successful businesses, marriages, and lives.

EXHIBIT 7
Relative Wealth Across Different Horizons
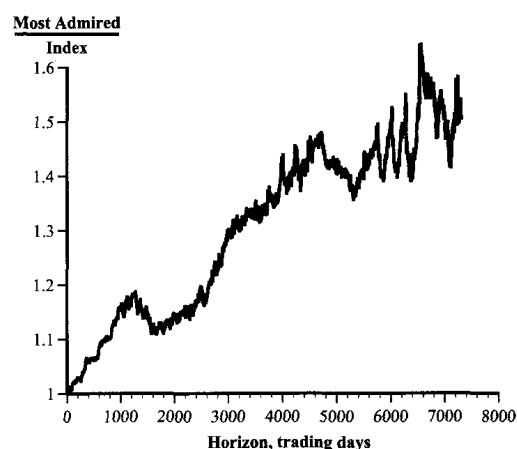
If we believe that "a bias for action" predicts success, a valid way to test the theory would be to identify companies that have a bias for action and companies that do not, and then see which companies do better over, for instance, a span of the next 10 years. The same holds true for secrets promising a successful marriage and a long life. Otherwise, we are merely staring at the past instead of predicting the future.

Looking backward, we will always find commonalities among already great companies. The interesting question is which characteristics will lead to future success. The future can seldom be seen by looking backward, and traits identified after selecting companies to analyze are unconvincing. The stock performance of the Great and Excellent companies after their identification in their respective best-selling books has been decidedly mediocre. These successful companies have not been remarkable investments, perhaps because their past successes were well known and their future regression to the mean were not fully anticipated. Their contrast with the

Most-Admired stocks is striking. The Most-Admired stocks performed better than the market, evidently because the *Fortune* survey captured intangible factors that investors had not fully taken into account. Information regarding which companies possessed these traits was useful because the traits were identified in advance, before the companies were selected.

REFERENCES

Anderson, J., and G. Smith. "A Great Company Can Be a Great Investment." *Financial Analysts Journal*, Vol. 62, No. 4 (2006), pp. 86-93.

Anginer, D., and M. Statman. "Stocks of Admired and Spurned Companies." *The Journal of Portfolio Management*, Vol. 36, No. 3 (2010), pp. 71-77.

Antunovich, P., D. Laster, and S. Mitnick. "Are High-Quality Firms also High-Quality Investments?" *Current Issues in Economics and Finance*, Vol. 6, No. 1 (2000), pp. 1-6.

Bannister, B.B. "In Search of Excellence: A Portfolio Management Perspective." *Financial Analysts Journal*, Vol. 46, No. 2 (1990), pp. 68-71.

Banz, R. "The Relationship Between Return and Market Value of Common Stocks." *Journal of Financial Economics*, 9 (1981), pp. 3-18.

Baumol, W.J., S.A.B. Blackman, and E.N. Wolff. *Productivity and American Leadership: The Long View*, Cambridge and London: MIT Press, 1989.

Chan, K. "On the Contrarian Investment Strategy." *Journal of Business*, 61 (1988), pp. 147-163.

Clayman, M. "In Search of Excellence: The Investorce: Viewpoint." *Financial Analysts Journal*, Vol. 43, No. 3 (1987), pp. 54-63.

——. "Excellence: Revisited." *Financial Analysts Journal*, Vol. 50, No. 3, (1994), pp. 61-65.

Collins, J. *Good to Great: Why Some Companies Make the Leap... And Others Don't*. New York, NY: HarperCollins, 2001.

De Bondt, W.F.M., and R.H. Thaler. "Does the Stock Market Overreact?" *Journal of Finance*, 40 (1985), pp. 793-805.

——. "Further Evidence on Investor Overreaction and Stock Market Seasonality." *Journal of Finance*, 42 (1987), pp. 557-581.

Fama, E.F. "Efficient Capital Markets: II." *Journal of Finance*, 46 (1991), pp. 1575-1617.

Fama, E.F., and K.R. French. "Dividend Yields and Expected Stock Returns." *Journal of Financial Economics*, Vol. 22, No. 1 (1988), pp. 3-25.

——. "The Cross-Section of Expected Stock Returns." *Journal of Finance*, 47 (1992), pp. 427-465.

——. "Size and Book-to-Market Factors in Earnings and Returns." *Journal of Finance*, 50 (1995), pp. 131-155.

——. "Forecasting Profitability and Earnings." *Journal of Business*, Vol. 73, No. 2 (2000), pp. 161-175.

Fortune. "Most Admired 2015." Fortune Magazine, 2015, fortune.com/worlds-most-admired-companies/.

French, K.R. "Data Library." 2013, mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Friedman, M. "Do Old Fallacies Ever Die?" *Journal of Economic Literature*, Vol. 30, No. 4 (1992), pp. 2129-2132.

Gardner, D., and T. Gardner. *The Motley Fool Investment Guide: How the Fools Beat Wall Steet's Wise Men and How You Can Too*. New York: Simon and Schuster, 1996.

——. "Farewell, Foolish Four." The Motley Fool, 2000, www.fool.com/ddow/2000/ddow001211.htm.

Jegadeesh, N., and S. Titman. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *Journal of Finance*, 48 (1993), pp. 65-91.

Kahneman, D., and A. Tversky. "On the Psychology of Prediction." *Psychological Review*, 80 (1973), pp. 237-251.

Keil, M., G. Smith, and M.H. Smith. "Shrunken Earnings Predictions Are Better Predictions." *Applied Financial Economics*, 14 (2004), pp. 937-943.

Keynes, J.M. *The General Theory of Employment, Interest and Money*. New York: Macmillan, 1936.

Kim, M., C. Nelson, and R. Startz. "Mean Reversion in Stock Prices? A Reappraisal of the Empirical Evidence." *Review of Economic Studies*, 58 (1991), pp. 515-528.

Lakonishok, J., A. Shliefer, and R.W. Vishny. "Contrarian Investment, Extrapolation, and Risk." *Journal of Finance*, Vol. 49, No. 5 (1994), pp. 1541-1578.

La Porta, R. "Expectations and the Cross-Section of Stock Returns." *Journal of Finance*, 51 (1996), pp. 1715-1742.

McQueen, G. "Long-Horizon Mean-Reverting Stock Prices Revisited." *Journal of Financial and Quantitative Analysis*, 27 (1992), pp. 1-18.

McQueen, G., and S. Thorley. "Mining Fool's Gold." *Financial Analysts Journal*, Vol. 55, No. 2 (1999), pp. 61-72.

Niendorf, B., and K. Beck. "Good to Great, or Just Good?" *Academy of Management Perspectives*, Vol. 22, No. 4 (2008), pp. 13-20.

Peters, T., and R. Waterman, Jr. *In Search of Excellence.* New York: Harper & Row, 1982.

Poterba, J.M., and L.H. Summers. "Mean Reversion in Stock Prices: Evidence and Implications." *Journal of Financial Economics*, Vol. 22, No. 1 (1988), pp. 27-59.

Reinganum, M. "Misspecification of Capital Asset Pricing: Empirical Anomalies Based on Earnings' Yields and Market Values." *Journal of Financial Economics*, 9 (1981), pp. 19-46.

Resnick, B.G., and T.L. Smunt. "From Good to Great to..." *Academy of Management Perspectives*, Vol. 22, No. 4 (2008), pp. 6-12.

Richardson, M. "Temporary Components of Stock Prices: A Skeptic Beat W." *Journal of Business & Economic Statistics*, Vol. 11, No. 2 (1993), pp. 199-207.

Rosenberg, B., K. Reid, and R. Lanstein. "Persuasive Evidence of Market Inefficiency." *The Journal of Portfolio Management*, 11 (1985), pp. 9-17.

Secrist, H. *The Triumph of Mediocrity in Business.* Evanston, IL: Northwestern University Press, 1933.

Sheard, R. "The Daily Dow." The Motley Fool, December 1997, www.fool.com/DDow/1997/DDow971230.htm.

Shiller, R.J. "Stock Prices and Social Dynamics." *Brookings Papers on Economic Activity*, 2 (1984), pp. 457-498.

Summers, L.H. "Does the Stock Market Rationality Reflect Fundamental Values?" *Journal of Finance*, 41 (1986), pp. 591-601.

Williamson, J.G. "Productivity and American Leadership: A Review Article." *Journal of Economic Literature*, Vol. 29, No. 1 (1991), pp. 51-68.

To order reprints of this article, please contact Dewey Palmieri at dpalmieri@iijournals.com or 212-224-3675.