21 APRIL 2020

Critical Thinking

PROFESSOR ALEX EDMANS

Confirmation Bias: Why We Might Not Think Critically

a. "Equality is better for everyone"
b. "CEO remuneration packages actively discourage innovation in UK's top companies"
c. "The outperformance of ESG strategies is beyond doubt"
d. "Professors on corporate boards increase profits"

These four statements have three things in common:

1. Those making them argue that they're backed up by research
2. Most people would like these statements to be true. Indeed, I have a particular interest in them being true:
    a. In last year's lecture, "Executive Pay: What's Right, What's Wrong, and What's Fixable", I argued that shares should be given to all employees, not just the CEO, so that the benefits of corporate success are distributed fairly
    b. The same lecture argued that executive pay should be reformed to encourage long-term thinking
    c. My inaugural lecture, "Purposeful Business: The Evidence and the Implementation", argued that companies that perform well on Environmental, Social, and Governance dimensions may also deliver higher returns t shareholders
    d. If this statement were true, I would be more likely to be offered board seats.
3. None of the statements is actually backed up by research.

Importantly, (2) and (3) are linked. Since most people would like these statements to be true, they accept these statements uncritically, even if the underlying evidence is flimsy. This is an example of confirmation bias: accepting a statement uncritically if it confirms what we'd like to be true, and rejecting a statement that contradicts it.

This lecture is on how we can think critically and avoid succumbing to confirmation bias. Note that "critical thinking" includes other topics such as logical reasoning (e.g. deduction, induction, and abduction), but I will focus on confirmation bias since it is so pervasive. My TED talk "What to Trust in a Post-Truth World" also addresses confirmation bias; I've deliberately chosen to focus on different aspects in this lecture to avoid duplication. Thus, the reader interested in this topic should find the TED talk of additional value.

To illustrate my points, I need to use examples. These examples are *not* chosen in order to "bash" particular studies. Indeed, as mentioned above, I have a personal interest in these studies being true. If my goal was to bash a study, it would be easiest for me to choose a non-impactful (and thus likely low-quality) one. I have deliberately chosen studies that are highly influential – in some cases I used to quote them myself because of my own confirmation bias – to illustrate how easy it is to think uncritically.

The purpose of this talk is to be entirely constructive. Companies underperform, economies stagnate, and societies malfunction. If improving performance was as easy as claimed by some viral talks, influential studies, and best-selling books, this would not be the case. Instead, the underperformance may be due to following conclusions that aren't actually backed up by evidence. Any elements of the talk that may appear sharp are only to highlight the severity of confirmation bias, and thus the importance of critical thinking.

The Neurological Basis for Confirmation Bias

A 2016 study, published in *Scientific Reports*, featured human subjects with liberal political views.[1] The researchers read out either a political statement that the subjects agreed with (e.g. "The death penalty should be abolished") or a non-political statement (e.g. "The primary purpose of sleep is to rest the body and mind"). The researchers then read out contradictory evidence and measured the subjects' brain activity. There was no effect when non-political statements were contradicted. When political statements were contradicted, the amygdala of subjects' brains was triggered. This is the same part of the brain that is activated when a tiger attacks you and induces a "fight-or-flight" response. Thus, some people respond to their views being contradicted as if attacked by a tiger.

We'll now look at various forms in which confirmation bias may manifest, and then discuss practical tips to avoid such a bias.

1. **Statements Without Facts**

Many views are expressed in the form of simple statements. When making a statement, you have the incentive to make it as one-sided as possible. For example, Senator Bernie Sanders claimed that "Wall Street CEOs who helped destroy the economy, they don't get police records. They get raises in their salaries." Such an extreme quote will be widely shared by opponents of capitalism, since it illustrates capitalism's alleged ills in a very stark manner. Moreover, an extreme quote reduces the need to back it up with evidence, by giving the impression that the point is so obvious that no proof is needed. Indeed, extreme quotes are often made without referring to evidence at all. The CEOs of Bear Stearns and Lehman lost nearly $1 billion of personal wealth when their firms collapsed. But, due to confirmation bias, many people sharing the quote did not check the facts.

2. **Misattributed Statements**

Sometimes, statements are backed up by references. However, these references may not actually show what the statement claims that it shows. But, if the statement is consistent with what most people would like to be true, readers won't bother checking the reference.

For example, the UK House of Commons Select Committee Report on Executive Pay stated that "the evidence is at best ambiguous on the impact of individual CEOs on company performance", with a footnote referring to the evidence submitted by "Professor Alex Edmans" to the Executive Pay inquiry. However, nowhere did my evidence suggest this. The closest sentence I wrote on the "impact of individual CEOs" was "CEOs with high equity incentives outperform CEOs with low equity incentives by 4-10% per year, and the researchers do further tests to suggest that the results are causation rather than correlation", which is the opposite of the statement.

Similarly, Malcolm Gladwell's book *Outliers* famously introduced the 10,000 hour rule – if only if you spend 10,000 hours on any cognitively complex task, you'll become an expert. He quotes research Anders Ericsson as evidence. I must admit that I viewed this best-selling book as authoritative, and thus used to quote it in the orientation address I gave to LBS MBAs each year. It confirmed what I wanted to be true – you can do anything you put your mind to – so I didn't bother to look up the original study by Ericsson. But I've since done so, and it makes no such claim. Indeed, Ericsson's own book, *Peak*, has a section "The 10,000 hour rule isn't really a rule".

A related issue is that a statement may be quoted out of context. Former GE CEO Jack Welch is frequently cited as saying "Shareholder value is the dumbest idea in the world". Two points that we've already discussed apply here. First, the extremity of this statement should ring alarm bells – even if shareholder value is not the best idea, there are likely far worse ideas. Second, the statement is not backed up by evidence. While Welch was an extremely successful CEO, he did not conduct a large-scale study comparing companies which pursued shareholder value with otherwise identical companies which pursued a different objective. But there's a third point – this statement is quoted out of context. The full sentence is "On the face of it, shareholder value is the dumbest idea in the world", which has a very different meaning. Welch then goes onto clarify his point and explain why shareholder value is not as "dumb" an ideas as it may initially seem.

---

[1] Kaplan, Jonas T., Sarah I. Gimbel and Sam Harris (2016): "Neural Correlates of Maintaining One's Political Beliefs in the Face of Counterevidence." *Scientific Reports* 6, 39589. For a humorous cartoon summarising this study, and confirmation bias more generally, see https://theoatmeal.com/comics/believe.

3. The Narrative Fallacy

Let's say that neither problem (1) or (2) apply – we have facts that are true. For example, Apple used to have the "why" statement "Everything we do, we believe in challenging the status quo." And Apple has indeed been extremely successful.

The *narrative fallacy* is the temptation to see a set of facts and ascribe a cause-effect relationship to these facts, even if there's no true link. We have this temptation because we want to be able to make sense of the world, and form links between phenomena even if they're unrelated. Indeed, when something happens that's difficult to explain (e.g. a tragedy), we tell kids "everything happens for a reason."

Continuing the above example, Simon Sinek's best-selling book *Start With Why* claims that Apple's success is due to it being driven by its "why". I very much like Simon's TED talks and my own book is on the importance of purpose. However, it is impossible to prove that Apple's "why" drove its success. There could be very many other contributing factors, e.g. Steve Jobs's ideas or his network of relationships. However, the narrative that success was due to Apple's "why" is particularly appealing to readers as anyone can adopt a "why" - whereas not everyone can suddenly think of a novel idea or has a network of relationships.

Moreover, often interpretations are imposed after-the-fact, reverse-engineered to fit the facts. Indeed, as Jobs himself pointed out in his 2005 graduation speech at Stanford University "You can't connect the dots looking forward; you can only connect them looking backwards". With hindsight, it may appear that Apple started with "Why", but it may have not actually pursued this strategy at the time – just as a snooker player may unintentionally pot a ball when aiming for a different pocket.

A solution might be to study many companies. If we look across many successful companies and find that they all have a common characteristic (e.g. humble leadership), then isn't this proof that the success was due to this common characteristic? Indeed, this is the approach taken by many best-selling business books, such as *In Search of Excellence*, *Good to Great*, and *Built to Last*, which studied 43, 11, and 18 companies respectively. However, even ignoring issues of correlation vs. causation (which I'll get to shortly), "all successful companies have humble leadership" *does not mean* that "all companies with humble leadership are successful". This is an error in logical reasoning. To show the latter, you can't just focus on companies that ended up successful – you'd have to study *all* companies with humble leadership (regardless of whether they turned out successful) and compare them with companies without humble leadership.

Indeed, force-fitting an explanation after the fact – whether to one or to many companies – is an example of spurious correlation. If there is no true relationship, then it will disappear in subsequent time periods (economists refer to this as "out-of-sample"). Indeed, many of the companies in the above books have subsequently underperformed, such as Xerox, Philip Morris, GE, Fannie Mae, Circuit City, Wells Fargo, and Ford.

As Nobel Laureate Daniel Kahneman points out in his book *Thinking, Fast and Slow*, this "regression to the mean" is to be expected if the true cause of the relationship is luck. He writes: "The basic message of *Built to Last* and other similar books is that good managerial practices can be identified and that good practices will be rewarded by good results. Both messages are overstated. The comparison of firms that have been more or less successful is to a significant extent a comparison between firms that have been more or less lucky. Knowing the importance of luck, you should be particularly suspicious when highly consistent patterns emerge from the comparison of successful and less successful firms. In the presence of randomness, regular patterns can only be mirages."

The standard way to address "luck" is to test for statistical significance. A significant test gives the likelihood that a statistical relationship arose from pure chance. The conventional significance level is 5% - we only conclude that a relationship exists if there is less than a 5% chance that the data could be due to luck. However, many studies have no tests of statistical significance.

4. Alternative Explanations

So if you find a statistically significant correlation between two variables (A and B), does this prove a relationship? For example, A is whether a company has a professor on the board, and B is performance. Does this mean that professors on the board cause superior performance? No, for two reasons:

- *Reverse causality.* B might have caused A. It's good performance that leads to professors being on the board. Professors might be risk-averse, hence preferring the ivory tower of academia to the cut-and-thrust of business. So they'll only join the board of a company that's already performing well.
- *Omitted variables.* There's a third variable, C, which causes both. For example, a forward-thinking board might want to increase its diversity of thought and appoint a professor. Thus, C causes A. Also, a forward-thinking board causes the company to perform well due to injecting it with good ideas. Thus, C causes B as well; there's no relationship in either direction between A and B.

Now the phrase "correlation does not imply causation" is very well-known. However, despite this, accepting correlation uncritically as causation is extremely common in business, particularly since people like to believe that there are certain practices that will improve performance. Indeed, many studies do contain a few sentences acknowledging that correlation doesn't imply causation, but as a perfunctory disclaimer – the authors typically market the report (through press releases, launch events, and adverts) as showing causation. For example, statements such as "Equality is better for everyone" and "CEO remuneration packages actively discourage innovation in UK's top companies" imply causal relationships even though the studies contain brief caveats.

Sometimes authors will claim to have separated causation from correlation. There are indeed valid methods of doing so, but often these claims are unwarranted. For further detail, see the article "A Layman's Guide to Separating Causation from Correlation … And Noticing When Claims of Causality are Invalid" (www.alexedmans.com/correlation).

To illustrate how correlation is often interpreted as causality, despite the "correlation does not imply causation" phrase being so well-known, consider the book *The Spirit Level: Why Equality is Better for Everyone*. I have chosen this example as the book has been extremely successful and influential, and I commend it for injecting the importance of inequality into the public debate; this discussion will be purely on its evidence base. The book gives scatter plots of country-level inequality against various country-level outcomes (e.g. obesity, happiness, health and social problems) and argues that there is a relationship to the naked eye. In addition to problem (3) (there are no tests of statistical significance), there are also no control variables. There are very many other country-level factors that could cause the outcomes, the most obvious being poverty. Countries with high inequality also have low average incomes, and it could be poverty, rather than inequality, which leads to these negative outcomes. The authors address this explanation by drawing a scatter plot of health and social problems against poverty and arguing there is no relationship to the naked eye. However, basic statistics tells you that it's invalid to rule out a "poverty" explanation by doing *separate* regressions on poverty and inequality. Instead, you need to do a *single* regression on inequality controlling for poverty (and for the myriad of other factors that may drive obesity, e.g. national diet, amount of PE taught at schools, availability of fitness facilities). Just as seriously, the authors forget about poverty explanation for the rest of the book; almost all their future graphs (with different outcome variables to "health and social problems") study inequality only.

Reverse causality is also an issue. It could be that inequality is the result of the various "outcomes", not the cause. For example, ill health may cause people to be unable to study or work, in turn leading to inequality. The authors acknowledge on rare occasions that correlation is not causation but they "believe" that this relationship is causal. Their arguments for this belief are invalid. One is that the strength of the relationship suggests that it's causal, but a relationship could be strong yet be in the other direction (or there could be a strong omitted variable). Another is the fact that a relationship has been found in many settings. Again, causality could be in the other direction in all of these settings, or an omitted variable could be present in all of these settings (and it need not be the same omitted variable in each one).

5. Universality Bias

This is phrase that I am introducing to describe our preference for "universal" explanations that apply elsewhere. We like to have a theory of everything. This allows us to make sense of the world more easily; we can also write articles or give talks on "Five ways to …" which apply to every company. This is related to the problem of the "narrative fallacy", but that discussion surrounded *internal validity*. Within a particular dataset (e.g. a group of companies and measures of their performance and CEO humility), we might ascribe a relationship when there is none. This discussion concerns *external validity*. Even if we could convincingly nail a relationship with a dataset, it may not apply to companies outside that dataset.

For example, even if we could prove that Apple's "why" caused its success, a "why" might not help other companies. Perhaps those in low-margin, hyper-competitive industries should focus more on cost control rather than finding their "why". This illustrates the danger of overextrapolating from single stories. Stories are vivid, they're powerful, they bring a topic to life and get re-told. They've thus been used successfully in business schools (as case studies), books, and TED talks. But stories tell you little unless they're backed up by large-scale data, because you can always hand-pick a story to support any viewpoint. Supporters of an exclusive focus on profit might use the story of GE under Jack Welch to show it can succeed. Opponents might use the story of Enron to show it can fail. Indeed, both GE and Enron are major business school case studies, but neither story tells us whether running a company for profit works in general.

Even if you show a relationship in a large sample of data, it may not apply outside that sample. Gladwell claimed that "The 10,000 hours rule says that if you look at any kind of cognitively complex field, from playing chess to being a neurosurgeon, we see this incredibly consistent pattern that you cannot be good at that unless you practice for 10,000 hours". However, the Anders Ericsson study was on violinists. He and his coauthors did not consider chess, neurosurgery, or other cognitively complex activities (nor did they claim a 10,000-hour rule exists even for violinists.). Limited external validity applies even to my own work. A paper I published in 2011 showed that US companies with high employee satisfaction outperform their peers.[2] Since then I've extended the study to non-US countries and found that the relationship only holds in those with flexible labour markets; it doesn't hold in countries with regulated labour markets.[3]

The above concerns the application of a result to fields that you haven't even analysed (e.g. neurosurgery). A related manifestation of the universality bias does involve analysing data in different fields, but imposing your "theory of everything" on it and not considering alternative explanations – as captured by the *Alice in Wonderland* phrase "sentence first, verdict afterwards". If authors have a hammer, everything looks like a nail. For example, the blurb for *The Spirit Level* reads: "Why do we mistrust people more in the UK than in Japan? Why do Americans have higher rates of teenage pregnancy than the French? What makes the Swedish thinner than the Greeks? The answer: inequality." But there may be very many other explanations for those phenomena, and the authors don't consider or test them.

6. Claims Not Supported by the Research

Even if we cast aside issues of internal validity (correlation vs. causation) and external validity, researchers may make claims that the study does not actually support. For example, a TED talk used to be titled "Want a more innovative company? Hire more women." The study correlated the share of revenues that stems from new products with a composite of 6 measures of diversity, including not only gender but other factors such as country of origin and age. Ignoring causality, and assuming that "share of revenues from new products" is a good measure of innovation (which it may not be – this measure could be high simply because you've failed to maintain your existing products), this study does not even document a correlation between innovation and gender diversity – the correlation is with a composite measure of diversity. It could be that other components drove the relationship.[4]

[2] Edmans, Alex (2011): "Does the Stock Market Fully Value Intangibles? Employee Satisfaction and Equity Prices." *Journal of Financial Economics* 103, 621-640.
[3] Edmans, Alex (2018): "Employee Satisfaction, Labour Market Flexibility, and Stock Returns Around the World."
[4] The title of the talk has since been changed to "How Diversity Makes Teams More Innovative", which still claims a causal relationship.

Moreover, some studies may make claims without correlating anything at all. A study claimed that "CEO remuneration packages actively discourage innovation in UK's top companies" yet does not have any measures of innovation at all, so there are not even correlations.[5] It finds that UK companies' remuneration packages have certain features (e.g. earnings-based bonuses), which they believe to be detrimental to innovation. However, without showing that these features actually deter innovation, this conclusion is invalid – it is a case of "sentence first, verdict afterwards." The paper simply *assumes* that earnings-based bonuses deter innovation because innovation expenditure decreases earnings. But the effect of incentives on behaviour is extremely, extremely complicated, which is why there is an entire literature on it. Indeed:

- The same logic implies that earnings-per-share incentives will encourage share buybacks, because share buybacks increase earnings-per-share. But there is no evidence of this, either in the US[6], or in the UK study PwC and I conducted for the UK government.[7]
- Some incentives can have the opposite effect to what intuition would suggest – e.g. incentives can *deter* activity by undermining intrinsic motivation.
- Earnings-based incentives may encourage the company to "get its act together" and improve in many dimensions, including innovation. For example, by reducing injury frequency, Alcoa got its act together and improved in many other dimensions including productivity. Indeed, evidence shows that takeover threat (which might cause a company to focus on earnings to avoid takeover) increases innovation.[8]

Practical Tips: What to Trust

I recognise the above examples paint a very bleak picture. They seem to suggest that we cannot trust anything – even statements in official government documents or in best-selling books. Even if a statement by an influential person may not be backed up by evidence (1). Even if it refers to evidence, the statement may be misquoting what the researchers actually claim (2). Even if the statement correctly conveys what the researchers claim, this claim may be invalid – the researchers may not have even studied the variable in question (6), they may be ascribing a relationship where none exists (3), they may not have considered alternative explanations (4), or the results may be overextrapolated beyond the initial context (5).

Unfortunately, the picture is indeed bleak. Very influential studies and books have a much weaker evidence basis than they claim. Given the neurological basis for confirmation bias, there is a huge incentive to convey results that others would like to be true. Your policy will get widespread support, your book will become a best-seller, and your report, blog, quote, or talk will be widely shared.

So, what are practical tips we can use to ensure we think critically and overcome confirmation bias?

1. For a statement, check if it is backed up by a reference. If it is a quote, check the context.
2. If it is, skim the reference and see whether it actually claims what the statement states that it claims. (Even if the reference is an academic paper behind a paywall, the abstract will always be available free of change. Moreover, the full version of most social science papers is available on www.ssrn.com).
3. If the authors' claim agrees with the statement, check that the authors' actual analysis supports the claims. Think whether there might be plausible alternative explanations, such as omitted variables, reverse causality, or luck.

Point (3) might be difficult for non-academics to verify, particularly if the paper is technical. So, the reader should check if the paper is published in a top peer-reviewed academic journal. Practitioner studies have significant value. They often have better access to data than academics, and are often a superior source for statistics. But academics have particular expertise in drawing relationships *between* statistics – teasing apart causation from correlation and

---

[5] "The Invisible Drag on UK R&D: How Corporate Incentives Within the FTSE 350 Inhibit Innovation"

[6] Bennett, Benjamin, J. Carr Bettis, Radhakrishnan Gopalan and Todd Milbourn (2017): "Compensation Goals and Firm Performance." *Journal of Financial Economics* 124, 307–330.

[7] Atanassov, Julian (2013): "Do Hostile Takeovers Stifle Innovation? Evidence from Antitakeover Legislation and Corporate Patenting." *Journal of Finance* 68, 1097–1131.

[8] "Share Repurchases, Executive Pay and Investment"

addressing alternative explanations. Importantly, academic studies have to undergo rigorous peer review to check their scientific accuracy. The very top journals have the highest standards, using the world's leading specialists to scrutinise a manuscript, and reject up to 95% of manuscripts. The 5% not rejected aren't immediately accepted either; instead, their status is "revise-and-resubmit". The reviewers highlight concerns that the authors need to address, and the paper can still be rejected at the next round. It's not unusual for a paper to take five years to be published after its first draft. A hard slog for the authors, but it helps ensure that the published results are correct.

Here's an example of the importance of evidence quality, to show it's not just an "academic" issue but can transform a study's implications for real-world practice. In the UK House of Commons' 2016 inquiry into corporate governance, a witness quoted research which "found that firm productivity is negatively correlated with pay disparity between top executive and lower level employees", referencing a January 2010 work-in-progress draft. The finished version had actually been published three years prior to the inquiry.[9] Having gone through peer review and tightened up its methodology, it found completely the opposite result:

- *We do not find a negative relation between relative pay and employee productivity.*
- *We find that firm value and operating performance both increase with relative pay.*

I also gave evidence at the inquiry, and I submitted follow-up evidence explaining that the study's results had been completely overturned in peer review. Despite publishing my follow-up evidence, the final report of the inquiry still wrote that "The [witness] states that "There is clear academic evidence that high wage disparities within companies harm productivity and company performance." " This is an example of why the picture is indeed bleak, if even House of Commons reports refer to evidence known to be tainted. Thus, it is essential that we exercise critical thinking; we cannot even take the conclusions of official reports at face value.

I stress the importance of a paper being published in a *top* peer-reviewed journal. That a journal calls itself "peer-reviewed" is far from sufficient to guarantee its rigour, since there's a vast range in the quality of reviewing standards. The analytics company Cabell's has a blacklist of 8,700 journals that claim to be peer-reviewed but actually aren't. Journal quality can easily be checked by looking at one of the freely-available lists of the best ones, such as the *Financial Times* Top 50. One doesn't need to be an academic insider to do this. Peer review isn't perfect – mistakes are made. Sometimes sloppy papers get accepted and good papers are rejected. But it's better to go with something checked than something unchecked.

Of course, every paper starts out unpublished. We can gauge the quality of a new paper by considering the authors, such as:

1. Their track record of top-tier publications *in the relevant field.* This helps avoid the issue of "halo effects", where a person is seen as a guru due to expertise in one field and thus is trusted as an authority in different fields. For example, because Jack Welch was a successful CEO at one firm, he is viewed as an authority on the effect of a shareholder value on performance on firms in general.
2. Do they have a hammer? Are they known for a "theory of everything" which may have blinded them to alternative explanations?
3. Relatedly, if they had found the opposite relationship, or no relationship, would they have still released the study? Or are their incentives strongly skewed to finding the relationship they are claiming?

Is it Practical to Check Every Source?

Are these practical tips indeed practical? Even if we don't read the actual analysis in every paper referenced, it takes time to check the quality of the journal it's published in, or the credentials of the authors. So we should only scrutinise the references for the most important points that an article makes. Moreover, we should be particularly discerning if:

---

[9] Faleye, Olubunmi, Ebru Reis and Anand Venkateswaran (2013): "The Determinants and Effects of CEO-Employee Pay Ratios." *Journal of Banking and Finance* 37, 3258–3272.

1. The article or research draws conclusions that you would like to be true (your own confirmation bias)
2. The article or research draws conclusions that the author would like to be true (the authors' or researchers' confirmation bias).
3. The article of research is one-sided, claiming "clear evidence" or universality. There are nearly two sides to almost every social science issue, and claiming that an issue is unambiguous suggests that the authors may not have seriously considered the other side. For example, the TED talk on diversity claimed "The answer was a clear yes, no ifs, no buts. The data in our sample showed that more diverse companies are simply more innovative, period." *The Spirit Level*'s subtitle is "Equality is better for everyone." An article claimed "The outperformance of ESG strategies is beyond doubt" when meta-analyses (studies of studies) find that the evidence is mixed at best.[10]

An Example: Test Your Skills

To test your critical thinking skills, let's consider the following opening paragraph of a *Harvard Business Review* article.

"Joe Bower and Lynn Paine "had me at hello" (to quote *Jerry Maguire*) with their new HBR article, "The Error at the Heart of Corporate Leadership." Laying out their data, they find that long-term oriented companies create more financial value and more jobs. In fact, if more American companies were focused on the long term, they estimate, investors would have an additional $1 trillion, workers would have an additional 5 million jobs, and the country would have more than an additional $1 trillion in GDP."

Should we take this article at face value?

1. Consider the author. Here, we don't even need to examine his credentials. The fact that the Bower-Paine article "had me at hello" is a strong signal of confirmation bias – he was willing to accept their conclusion uncritically.
2. Check the reference. It turns out that the Bower-Paine article does not lay out any data or make any such claim. There is, however, a separate article in the same *HBR* issue as Bower-Paine (by different authors) that makes the claim. It has various measures of long-term behaviour (e.g. investment) and finds that long-term companies subsequently perform better.
3. Check for alternative explanations (internal validity). Reverse causality is an issue. A basic finance principle tells you that when companies' future outlook is better, they should invest more. Omitted variables are an issue. It could be that a good CEO invests more (as she has better ideas) and also improves long-term performance.
4. Check for external validity. The study states "Extrapolating from the differences above, we estimate that public equity markets could have added more than $1 trillion in asset value … companies could have created five million more jobs in the United States – unlocking as much as $1 trillion in additional GDP." However, such an extrapolation to other companies beyond the sample is invalid. It will not make sense for all companies to invest more, e.g. those in declining industries. Moreover, if all companies within an industry make "arms-race"-type investments, such as some types of advertising, this split the pie (taking value or jobs away from competitors) rather than growing the pie and adding to aggregate welfare.

Some of this summary is adapted from my book, *Grow the Pie: How Great Companies Deliver Both Purpose and Profit* (Cambridge University Press).

---

[10] Renneboog, Luc, Jenke Ter Horst, and Chendi Zhang (2008): "Socially Responsible Investments: Institutional Aspects, Performance, and Investor Behavior." *Journal of Banking and Finance* 32, 1723–1742.