

Evaluating Research¹

Alex Edmans, aedmans@london.edu
Professor of Finance, London Business School
Academic Director, Centre for Corporate Governance
Managing Editor, *Review of Finance*

Writers, such as authors and journalists, often cite research-based studies to support a point. This is a welcome development in an increasingly fact-free world. However, a danger is that there is a vast range in the quality of research-based studies:

- Some make the most basic methodological mistakes. For example, some studies on the returns to investing in shares ignore dividends.
- Some interpret the data incorrectly. Most studies investigate the link between two factors (A and B), such as inequality and health, and find a *correlation* between the two – countries where inequality is high also have less healthy populations. They claim *causation* from A to B – inequality causes ill health, and so a country should tackle inequality in order to improve its health. However, as is well-known, *correlation does not imply causation*, for two reasons.
 - There may be *reverse causality* – the causation may be in the other direction. It may be that ill health causes inequality, as it prevents people from going to work. If so, the solution is to tackle health directly, rather than inequality.
 - There may be *omitted variables* that affect both A and B. For example, democratically-elected governments may both reduce inequality and improve health. If so, the solution is to improve democracy – the root cause of inequality – rather than inequality which is only a symptom.
- Some over-extrapolate from the results and make claims that the data does not actually support. For example, Malcolm Gladwell’s famous “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 study that his claim was based on only investigated violin players.² The results may not apply to chess or neurosurgery.

This matters. Readers will readily share an article that confirms their own view of the world, even if the research that it cites is flawed. In turn, this can lead to the spread of misinformation – indeed, the anti-vaccination movement was founded on flimsy research which claimed that vaccination causes autism, and the vandalism of 5G masts was sparked by unverified stories that they exacerbate coronavirus. In addition, newspapers have frequently had to publish apologies for misquoting research, or quoting low-quality research.

However, a key challenge is that writers do not have the time or specialist expertise to delve into the weeds of every paper and check its methodology. The purpose of this note is to provide a short user’s guide to discern whether you can trust research. This involves performing six simple and rapid checks:

¹ Most of this material is adapted from Chapter 10 of “Grow the Pie: How Great Companies Deliver Both Purpose and Profit” (Cambridge University Press).

² Ericsson, K. Andres, Ralf Th. Krampe, and Clemens Tesch-Romer (1993): “The Role of Deliberate Practice in the Acquisition of Expert Performance.” *Psychological Review* 100, 363-406.

1. Does the Research Actually Exist?

This may seem an obvious question, but there have been several cases where journalists have written articles based on an authors' press release, when the actual study does not yet exist. Moreover, the study should be in the public domain rather than for sale or only available from the authors upon request. Not putting the study in the public domain signals that the authors are unwilling for it to be subject to scrutiny.

Sometimes a study is indeed available, but an abridged version – it may only describe the results, but not the underlying methodology. For example, a study claiming that ethical companies perform better should explain how it measures whether a company is ethical. How we can trust this claim hinges critically on how reliable the ethicality measure is, so it is critical to make the methodology available for scrutiny.

Importantly, a writer only needs to check whether the study exists and see if there is a methodology section. S/he does not need to read the study in detail.

2. Is the Research Based on Actual Data?

This may also seem an obvious question, but several studies are not based on actual data but simply canvassing people's opinion. For example, a study may report that "70% of board members think CEOs do not improve company performance." This is very different to studying this question with data – studying whether CEOs actually improve performance.

Certainly, surveys are useful to learn people's opinion. However, they should only be reported as showing people's opinion, rather than showing what actually happens.

Relatedly, the data needs to be related to claims made by the authors. For example, one consultancy issued a press release on its study entitled "CEO remuneration packages actively discourage innovation in UK's top companies." However, the study did not contain any data on innovation. It gathered data on CEO remuneration packages, showed that they contain bonuses, and *assumed* that bonuses discourage innovation.

3. Is It Published in a Top Peer-Reviewed Journal?

Many studies are conducted by organisations such as professional service firms and think tanks. Such studies have significant value. They often have better access to data than academics and are often a superior source for statistics. However, academics have particular expertise in drawing relationships *between* statistics – teasing apart causation from correlation and addressing alternative explanations.

Importantly, academic studies have to undergo rigorous peer review to check their scientific accuracy. Neither practitioner studies nor books have to undergo such checks. Peer review addresses not only honest mistakes in execution, but also deliberate bias, such as a pharmaceuticals company funding a report on its own drugs, or an academic claiming to have uncovered a scandal to become famous. 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 are not immediately accepted either; instead, their status is

‘revise-and-resubmit’. The reviewers highlight concerns that the authors need to address (such as the three errors mentioned above), and the paper can still be rejected at the next round.

Academic peer review matters; it is not just a rubber stamp. Indeed, the results of a study can be completely overturned by the peer review process. In the UK House of Commons’ 2016 inquiry into corporate governance, a witness quoted evidence 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 in 2013.³ Having gone through peer review and tightened up its methodology, it found 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.”

The stringency of the peer review process is critical. That a journal calls itself “peer-reviewed” is far from sufficient, since there is a vast range in the quality of reviewing standards. The analytics company Cabell’s has a blacklist of 8,700 journals that it classifies as “predatory”, for example claiming to be peer-reviewed when they actually are not.

Importantly, journal quality can easily be checked by looking at one of the freely available lists of the best ones. For business, there is the list of the *Financial Times* Top 50 journals; for science, the National Institute of Environmental Health Sciences and Scimago also produce lists. For other fields, the Impact Factor of a journal can be easily looked up and is a measure of quality.

Peer review is not perfect – mistakes are made. Sometimes sloppy papers get accepted and good papers are rejected. However, it is better to go with something checked than something unchecked. Some writers claim “publication bias” as a licence to ignore whether a paper’s been published and quote whatever study they like. Their charge is that journals only publish papers that support traditional orthodoxies. This is not how the publication process works. As journal editors, we wish to publish new papers that change the way people think. A journal’s Impact Factor is the number of times its articles are cited. The first paper in any new area will be hugely cited; the tenth in a well-established field will not.

One caveat to this question is that a small number of practitioner studies may be of sufficient quality to be published in an academic journal, but do not seek such an outcome as academic publication is not their objective. An example is studies by economics consultancies (or the economics department of a professional services firm) commissioned by the government or a regulator. It is critical to distinguish between studies undertaken to *inform* and studies undertaken to *advertise* the organisation’s brand or services. To do so, an important question to ask is: *would the organisation have published the study if it had found the opposite result?* For example, many consultancies will publish studies showing that ethical companies perform better (since making this claim is good for their brand) but not if they found the opposite. These incentives are particularly strong if the study uses the consultancy’s own measure of ethicality, since it shows that their measure “works”. Even if the answer is “No”, this does not automatically mean that the paper is wrong, but it must be approached with caution.

³ 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.

4. What are the Credentials of the Authors?

Of course, every academic paper starts out unpublished. How do we gauge the quality of a new paper? The fourth dimension to check is the credentials of the authors. One relevant factor is the quality of their institution, which we can compare against freely available lists of the top universities. This is not elitism, but simply a desire to use the best evidence. We would listen more closely to a medical opinion from the Royal Marsden Hospital than one we have never heard of.

It is certainly not the case that studies by top institutions are always correct and those by others are always wrong. Thus, a second factor is the authors' track record of top-tier publications, which is easy to find as nearly all academics make their CVs available. Indeed, we carefully scrutinise the credentials of an expert witness in a trial. Again, this does not mean that well-published authors are always right. Credentials are simply one factor to assess when evaluating evidence, just as a company's brand is one consideration in a purchasing decision, or an undergraduate's university is one element in an entry-level hiring decision. A useful question to ask is the following: *If the same study was written by the same authors, with the same credentials and had the opposite results, would we still be willing to believe it?*

Importantly, it is critical to scrutinise whether the authors have credentials *in the relevant field*. This helps avoid the issue of "halo effects", where a person with expertise in one field is seen as a guru in different fields. Liverpool FC manager Jürgen Klopp honourably declined to give his opinion on coronavirus, fearing that people might believe him because he is famous. While it might seem obvious that you should not believe a football manager on a scientific issue, actors have persuaded people to not vaccinate their children.

Less obviously, leading CEOs often give views on business. For example, former GE CEO Jack Welch is widely quoted for claiming that shareholder value is "the dumbest idea in the world." Welch was certainly an influential CEO at one firm, but has not investigated the effect of a shareholder value orientation on performance in firms in general – a question that warrants academic study rather than business leadership. Some doctors are proclaiming miracle cures for the coronavirus, even if their expertise is not in pathology or epidemiology. People ask me whether we should invest in the stock market after it has fallen due to the coronavirus, because I am a Professor of Finance. However, my expertise is in individual companies rather than macroeconomic conditions.

One warning sign is gurus who over-embellish their credentials with unverifiable claims, as this often signals that they lack genuine expertise. Common examples are as follows:

- "Best-selling author". There is no clear definition of this status – whether you need to be in the top 10, 100, 1000; whether for all books, or for books in a very small sub-field; or for how long (Amazon's best-seller list is updated every hour). See the Wikipedia article <https://en.wikipedia.org/wiki/Bestseller> for the ambiguity of this claim.
- "One of the world's top economists." Again, there is no clear ranking of economists. The claim is valid if a metric is cited (e.g. publications or citations).
- "International keynote speaker". This only requires the person to have given one keynote speech at one country outside the UK, even at a minor conference.
- "Award-winning professor." This only means that the professor has won one award, and says nothing about the competitiveness of the award.

5. Are There Alternative Explanations?

The fifth dimension to check is whether there are alternative explanations for the authors' results. One does not need to be an academic insider to conduct this check, because most alternative explanations are based on common sense rather than methodological technicalities. Readers can ask themselves whether the results could be driven by reverse causality or omitted variables. I commonly share academic papers on LinkedIn. If I post an article whose findings go against public opinion, there is no shortage of comments pointing out alternative explanations – so it is certainly feasible to think of them. But if the paper confirms current thinking, it is accepted at face value.

Importantly, it need not take long to figure out if there are alternative explanations. The introduction of an academic paper (typically 4-6 double-spaced pages) should be both fully self-contained and non-technical. It aims to give all the paper's punchlines – including how it deals with alternative explanations – without readers having to delve into the actual paper.

This question also can be asked of both published and unpublished papers. Most unpublished papers on social science are freely available at the Social Science Research Network, www.ssrn.com. Even if a paper is published in a journal that is behind a paywall, the pre-publication version typically remains available on SSRN.

6. Is It Balanced?

The final dimension to check is whether an article is balanced. In social sciences (e.g. business or economics), there are two sides to almost every issue, and it is difficult to definitively prove a result beyond doubt. Unlike in physical sciences, you cannot conduct controlled experiments where you change only one variable (e.g. inequality) and hold everything else constant (e.g. the quality of a country's government).

Writers should thus be wary of research that claims to have found “clear evidence” or “proof”, or shown something “beyond doubt”. Similar to overexaggerated author credentials, overexaggerated claims of proof often mask actual holes in the analysis. Researchers who claim “clear evidence” may not have seriously considered alternative explanations. Moreover, there is substantial incentive to claim “clear evidence” as a study is more likely to be widely shared than one with nuanced findings. In contrast, careful researchers will make it clear what their study can and cannot show.

The above concerns research that gathers its own primary data and conducts analyses. Many practitioner studies do not gather primary data, but aggregate the findings of existing academic papers. Again, it is rarely the case that the academic evidence is unambiguously in favour of one side. Any study claiming this has likely chosen to only include the academic evidence that supports what the study would like to show, and to deliberately exclude contradictory evidence. If a study claims that “There is clear evidence that X increases Y”, a simple check is to put “X decreases Y” into a search engine and see if there are links to high-quality evidence showing the opposite.

In sum, a balanced study is more likely to be written with the intent to inform and contribute to knowledge, rather than to make the authors famous or boost the organisation's brand.