

The Business Case for Diversity: A Critical Look at the Evidence

Professor Alex Edmans, London Business School

www.alexedmans.com, aedmans@london.edu

This is taken from my 27/2/18 response to the Financial Reporting Council's consultation on the Corporate Governance Code. The full response is at <http://bit.ly/AlexCGC>.

Paragraph 58 of the consultation reads:

There is clear evidence that greater female representation in the boardroom and senior management has a positive impact on performance. More recently, research has found a statistically significant relationship between ethnically and gender diverse leadership teams and better financial performance. Companies that focus on increasing diversity in the boardroom, in their executive teams and across their workforces as a whole can expect a positive impact on their performance.

I am a very strong supporter of gender diversity, ethnic diversity, and more broadly diversity in all forms (in particular, diversity of thinking in addition to diversity on observable dimensions). I took over as Managing Editor of the *Review of Finance*, the #1 academic finance journal in Europe, in 2017 and later that year, at the first available opportunity, I appointed the first female Editors onto the board in our 21-year history. However, I did this not because there was any evidence that gender diversity in editorial boards improves performance. I did this because (a) I thought it was the right thing to do. Diversity is intrinsically desirable in its own right, rather than an instrumental way to improve performance; (b) I thought that the two women I appointed are excellent in their own right, rather than appointing them to fulfil a quota or tick the box.

Problems With The Cited Evidence

My concern with paragraph 58 is that the evidence base for diversity improving performance is much weaker than stated (again, I stress that this does *not at all* mean that diversity is not desirable. Diversity is desirable in its own right, not as an instrumental way to improve performance). The first sentence refers to “clear evidence” that diversity has a “positive impact on performance”. For the “clear evidence”, the paper quotes one single McKinsey study. It is almost always possible to find a study that supports *any* viewpoint. There are studies showing that vaccination is beneficial, others showing that vaccination is harmful. Moreover, there is substantial variation in the quality of studies – some make very basic errors, confuse correlation with causation, or fail to consider alternative explanations. As a result, one should pay particular attention to studies published in the world's best peer-reviewed journals. Peer review involves a paper being independently scrutinised by the world's leading experts to verify the rigour of the methodology and whether the claims are actually supported by the data. Often, the results of a paper can markedly change – and even switch direction – after going through peer review. At the *Review of Finance*, I reject 97% of papers, so the peer review process is very tough. While peer review is not perfect (mistakes can sometimes be made), a verified study is more reliable than an unverified one. Indeed, much of what we learn about science and healthcare is from the very best academic studies, rather than from research by pharmaceutical companies.

Thus, a single non-peer-reviewed study does not constitute “clear evidence”. This problem is particularly a concern given confirmation bias – one would like to find evidence that supports one’s own view of the world. Being a big supporter of diversity, I would be delighted if there were clear evidence on its benefits, but this paper does not constitute it. For further discussion on the dangers of confirmation bias in the interpretation of evidence, and practical ways to address it, please see my TED talk “What to Trust in a Post-Truth World” at <http://bit.ly/trusted>.

In addition, the study does not show a “positive impact on performance”. That phrase implies *causation* rather than mere correlation. Causation is also implied in the final sentence of paragraph 58, which states that increasing diversity will *cause* higher performance. However, the McKinsey study is very clear that it only shows a correlation. In their executive summary on p1, they find that “where women are most strongly represented at board or top-management level are also the companies that perform best”. This is a statement of correlation, not causation. On p12, they explicitly state that their goal is “to examine whether greater gender diversity might correlate with better economic performance”, and even more explicitly on that same page they state that “correlation is not necessarily cause.”

Why might correlation not imply causation in this setting? It may be that a high-quality board understands the intrinsic value of diversity, and appoints women executives; in addition, a high-quality board improves firm performance anyway. Thus, there is a positive correlation between diversity and firm performance without causation. Or, causation could run the other way. Perhaps only once a company starts to perform well can it be forward-thinking about diversity and appoint women to senior roles (regrettably, given prior biases, the “default” or “safe” option might have been to appoint men).

Moreover, the evidence even for correlation is very weak. The evidence is given in Exhibit 9, suggesting that “companies with a higher proportion of women in their top management have better financial performance”, measured by average ROE, average EBIT, and stock price growth. However, there are several problems with these results:

- They compare diverse firms to the industry average. This misses very many control variables, such as firm size, age, growth opportunities, recent performance etc. Small firms and young firms typically have lower ROE and EBIT. A paper without basic controls could never be published in a peer-reviewed journal with even only modest standards.
- Performance is measured over only a two-year horizon. This is far too short to make meaningful inferences. The FRC correctly recognises in other areas of the pre-consultation that investors and companies must think long-term and not make inferences based on short-term stock price fluctuations. I fully agree, and so is difficult to put much weight on these results.
- The study does not even check for statistical significance. Given the very short time periods, the results could be driven by luck or outliers.
- The stock price growth measure ignores dividends, which is a very basic omission.
- The study uses inconsistent time periods. It uses 2003-5 for ROE and EBIT, and 2005-7 for stock price growth, making one wonder whether it selected the time periods that led to the most favourable results.

What Does Rigorous Evidence Show?

If the single paper cited by footnote 58 does not show that gender diversity improves firm performance, what are the findings of other studies? Given the importance of gender diversity, there have been a large number of studies on this question. An effective way to learn the overall consensus is to read a meta-analysis. This is when researchers summarise the findings of a multitude of individual studies to reach a general conclusion. Two recent meta-analyses have been undertaken on this topic: Post and Byron (2015) who synthesize 140 studies, with a combined sample of more than 90,000 firms from more than 30 countries. These include some unpublished studies (e.g. conference or working papers); Pletzer, Nikolova, Kedzior, and Voelpel (2015) have a more selective sample of 20 studies all published in peer-reviewed journals. The findings of these meta-analyses are summarised in an excellent synopsis (<http://knowledge.wharton.upenn.edu/article/will-gender-diversity-boards-really-boost-company-performance/>) by Katherine Klein, the Edward H. Bowman Professor of Management at Wharton and the Vice Dean of the Wharton Social Impact Initiative. She writes:

Research conducted by consulting firms and financial institutions is not as rigorous as peer-reviewed academic research. Here, I dig into the findings of rigorous, peer-reviewed studies of the relationship between board gender diversity and company performance. Spoiler alert: Rigorous, peer-reviewed studies suggest that companies do not perform better when they have women on the board. Nor do they perform worse. Depending on which meta-analysis you read, board gender diversity either has a very weak relationship with board performance or no relationship at all. ...

In sum, the research results suggest that there is no business case for — or against — appointing women to corporate boards. Women should be appointed to boards for reasons of gender equality, but not because gender diversity on boards leads to improvements in company performance.

Klein's article gives several logical reasons for why such a relationship may fail to exist, even though it may seem intuitive that diversity brings benefits such as differing viewpoints.

Moreover, most of the studies reviewed by the meta-analyses document only correlation, given the difficulty in showing causation. In addition, a potential disadvantage of a meta-analysis is that it weights all studies equally. However, as mentioned earlier, there is substantial variation in the quality of studies; even among published papers, there is a very wide range of standards set by journals. The most rigorous causal evidence on diversity of which I am aware is by Kenneth Ahern and Amy Dittmar (2012) and entitled "The Changing of the Boards: The Impact on Firm Valuation of Mandated Female Board Representation". It was published in the *Quarterly Journal of Economics*, one of the most elite journals in the world. It studies the impact of the 2003 Norway law mandating that 40% of directors be women. The methodology is extremely rigorous. It does not study what happened to firm performance after the law was passed, because this could be driven by other things going on in 2003 (e.g. macroeconomic conditions). Instead, it divides Norwegian firms into a "treated" group (those with no female directors, who are most affected by the law) and a "control" group (those with at least one female directors, who were less affected by the law). Comparing across these groups, they found that the quota led to a large decline in

firm valuation over several years; it also “led to younger and less experienced boards, increases in leverage and acquisitions, and deterioration in operating performance”.

Now, this study is far from the last word either. It is still only one study (albeit a very rigorous one). In addition, it does not show (or claim to show) that diversity is bad. It shows that quotas can be bad, because there might not be enough qualified people within the minority group – a quota that 40% of a board should be academics would likely be even more harmful, even if academics can add value to a board. However, the study is relevant because it shows how nuanced this issue is, and how there is not “clear evidence” in one direction. Moreover, it highlights caution in statements advocating that firms increase diversity (along any dimension) based on “evidence” that it improves financial performance. Increasing diversity to meet quotas or respond to regulatory pressure may worsen performance. Instead, boards should very carefully consider diversity as a valuable attribute of a potential board member or executive, and seriously take it into account alongside the other characteristics that it uses to guide appointments. In this sense, I endorse the FRC’s suggestions in Q9, Q10, and Q11 to encourage firms to voluntarily build up diversity in the boardroom, executive pipeline, and in the company as a whole – but would caution against going further and setting targets or implying that high diversity is “good” and low diversity is “bad”. The optimal level of diversity depends on many factors, including the pipeline for talent in a particular industry. More generally than the diversity issue, I hope that this response is helpful for other aspects of the consultation by highlighting the need to be very careful with evidence.

Diversity is highly desirable in its own right, and firms should pursue it even in the absence of a target and evidence showing that it instrumentally improves performance. It would be a sad world if the only reason firms increased diversity were to obtain higher performance or meet a regulatory target. Companies must give all of their workers equal opportunities, pay, and promotion prospects regardless of their gender, ethnicity, sexual orientation or other characteristics. This is simply the right thing to do.