

Divergent ESG ratings

Elroy Dimson, Paul Marsh, and Mike Staunton

This version: 19 July 2020

Elroy Dimson is a professor of finance and chairman of the Centre for Endowment Asset Management at the Judge Business School, Cambridge University, and a bye-fellow of Gonville and Caius College, Cambridge U.K.*

e.dimson@jbs.cam.ac.uk

Paul Marsh is an emeritus professor of finance at London Business School, London, U.K.

pmarsh@london.edu

Mike Staunton is director of the London Share Price Database at London Business School, London, U.K.

mstaunton@london.edu

* Corresponding author. Address: 2 Arlington, London N12 7JR, United Kingdom. Landline +44 20 8445 9988. Mobile +44 7966 908 212. Skype: elroydimson.

Divergent ESG ratings

Elroy Dimson, Paul Marsh, and Mike Staunton

This version: 19 July 2020

Abstract:

Responsible investors require data to underpin their stock and sector selections. Regardless of the rating agency, bond ratings for a particular issuer are broadly similar. This is not the case for ESG ratings. Companies with a high score from one rater often receive a middling or low score from another rater. This article examines the extent of, and reasons for, disagreement among the leading suppliers of ESG ratings. The weightings given to each pillar of an ESG rating also vary across agencies. Many asset managers contend that ESG ratings can help investors to select assets with superior financial prospects, and the authors therefore review the investment performance of portfolios and of indexes screened for their ESG credentials. In the authors' opinion, ESG ratings, used in isolation, are unlikely to make a material contribution to portfolio returns. (136 words)

Highlights:

- Data is essential for making investment decisions, and most institutions rely wholly or partly on external providers of ESG data. However, there is minimal correlation between ESG ratings from alternative agencies.
- This article explains why different raters' appraisals diverge. The article also examines whether ESG is associated with subsequent fund or index outperformance.
- Over the longest available period, ESG indexes did not outperform the "parent" index from which they are derived. Equally, however, for the two indexes that most closely represent what ESG investors actually do (FTSE4GOOD and MSCI-ESG) there is no evidence of underperformance.

Topics: Equity portfolio management, ESG investing, foundations and endowments, performance measurement, wealth management

JEL codes: G11 (*Portfolio Choice, Investment Decisions*), G30 (*Corporate Finance and Governance: General*), G41 (*Role and Effects of Psychological, Emotional, Social, and Cognitive Factors on Decision Making in Financial Markets*), M14 (*Corporate Culture, Diversity, Social Responsibility*)

It is increasingly important for asset owners to explain how they want their portfolios configured from an ESG perspective. Without data, investors cannot implement their strategies. If they are simply excluding stocks in certain sectors, they may get by with industry codes and a modicum of supplementary information. But for more sophisticated exclusionary screening, for positive screening, and particularly for ESG integration, they need detailed stock-by-stock analysis and ratings. For many, ESG rating services are the backbone of responsible investing.

The first ESG rating agency was the Ethical Investment Research and Information Services (EIRIS) agency launched in the U.K. in 1983. Since then, and especially in recent years, the ratings industry has expanded enormously. The global market for ratings is projected to grow by 2024 to annual revenues of \$500 million (Nauman 2019).

Organizations providing ESG rankings include major index companies such as MSCI and FTSE Russell; standalone providers, some offering a full-range service and others focusing on specialist niches such as emissions; rating agencies, such as Moody's and S&P (who are also index providers); and financial data companies, such as Refinitiv, Morningstar (the owner of Sustainalytics) and FactSet. As Walter (2019) explains, the industry has seen a wave of consolidation with well-known providers often changing ownership more than once.

Raters can base their assessment on public information, though many send questionnaires, ask to hold an interview or make site visits. Larger companies can afford to participate in these increasingly numerous and frequent appraisals; smaller ones cannot.

In this article, we discuss the important role of ESG metrics and the disconcerting lack of agreement across raters, review explanations for divergent scores, examine whether ESG ratings are predictive of investment returns, and discuss the implications for ESG investors.

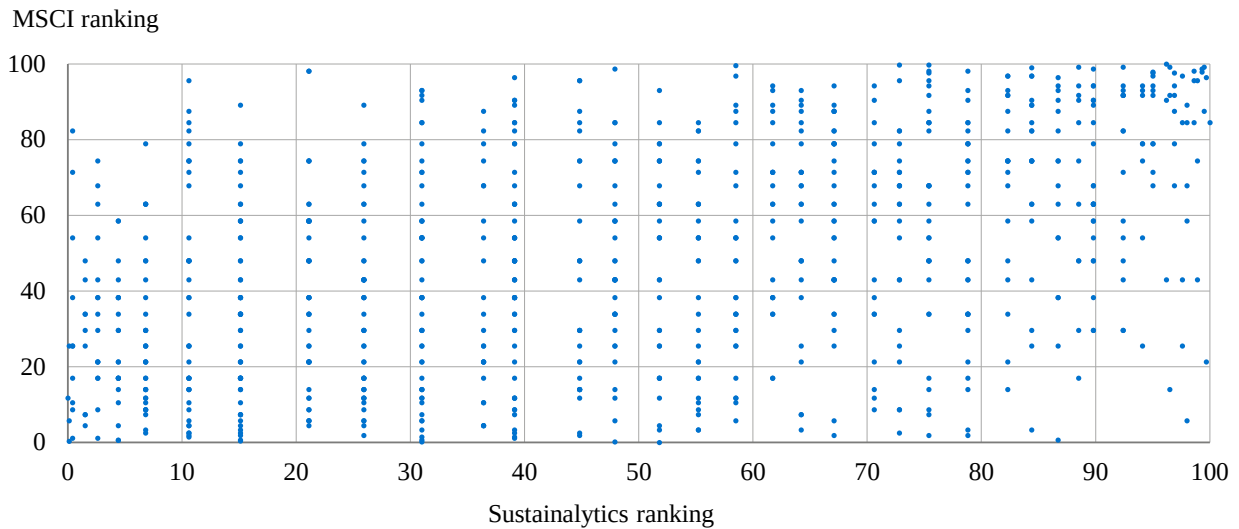
DIVERGENT RATINGS

An important issue is “convergent validity,” or the extent to which different raters agree. Evaluations that purport to measure the same variable should generate positively correlated scores. Not only should the score correlate with related variables but it should not correlate with dissimilar, unrelated ones. Different raters may, of course, generate diverse scores because they are measuring different aspects of ESG behavior. To illustrate, we would not necessarily expect a high correlation between rankings of the Top Twenty universities, if one rater were measuring research, another teaching, and a third sporting success. Similarly, users of ESG ratings may differ in the dimensions they value.

A widely cited example was the ESG score of the world's most valuable carmaker: in 2018-19 Tesla held an exemplary “AA” rating from MSCI, whereas FTSE ranked the company it very low and Sustainalytics put the firm in the middle. The discrepancy reflected the fact that MSCI judged the product to be almost perfect on carbon emissions, in contrast to FTSE which focused on factory emissions and regarded the firm as a serious offender (Dimson, Marsh, and Staunton 2020).

One might expect that the frequent ownership changes of leading raters, coupled with fierce competition, would lead to convergence in ratings. However, although the mainstream rating agencies are regarded as potential substitutes by many users, they have a remarkably low level of agreement. **Exhibit 1** shows a scatter plot of the overall ratings for 878 U.S. companies from two leading raters. There is a barely perceptible relationship, and the overarching impression is one of substantial disagreement.

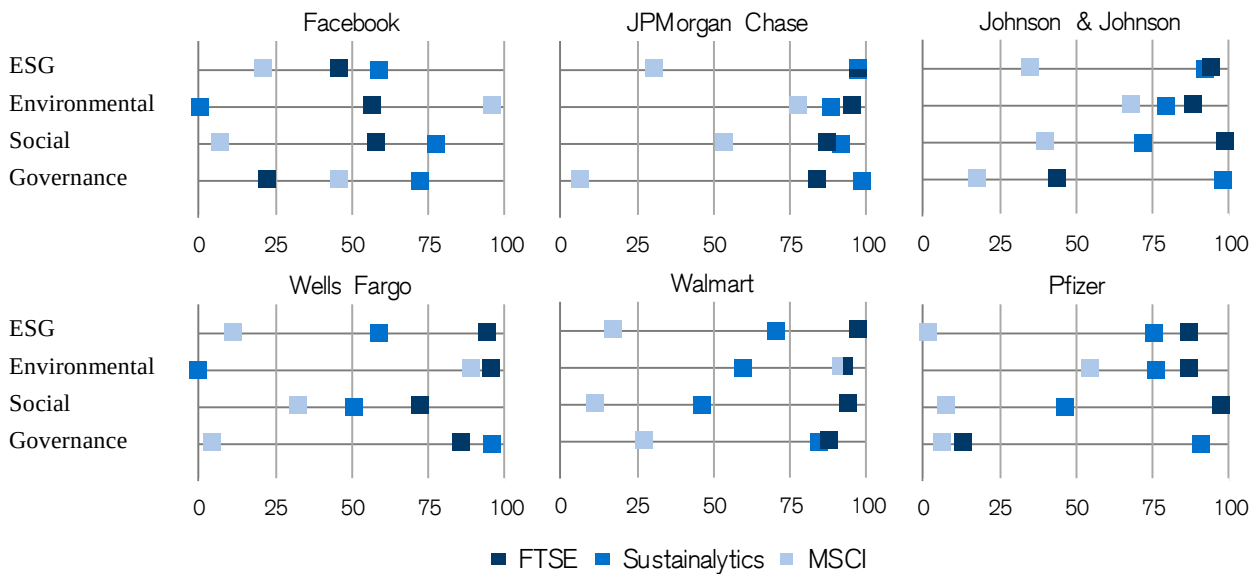
Exhibit 1: MSCI vs. Sustainalytics rankings at start-2019



Source: MSCI and Sustainalytics ranks for 878 US companies; analysis by the authors.

To emphasize the differences, we examine the ESG ratings provided by three providers (FTSE Russell, Sustainalytics, and MSCI) for some large, well-known companies: Facebook, JP Morgan Chase, Johnson & Johnson, Wells Fargo, Walmart and Pfizer. We present in **Exhibit 2** each company's ratings for the environmental, social and governance pillars taken separately, as well the overall ESG score. The divergences across raters can be quite stark.

Exhibit 2: Divergence in ratings across large, U.S. companies, 2019



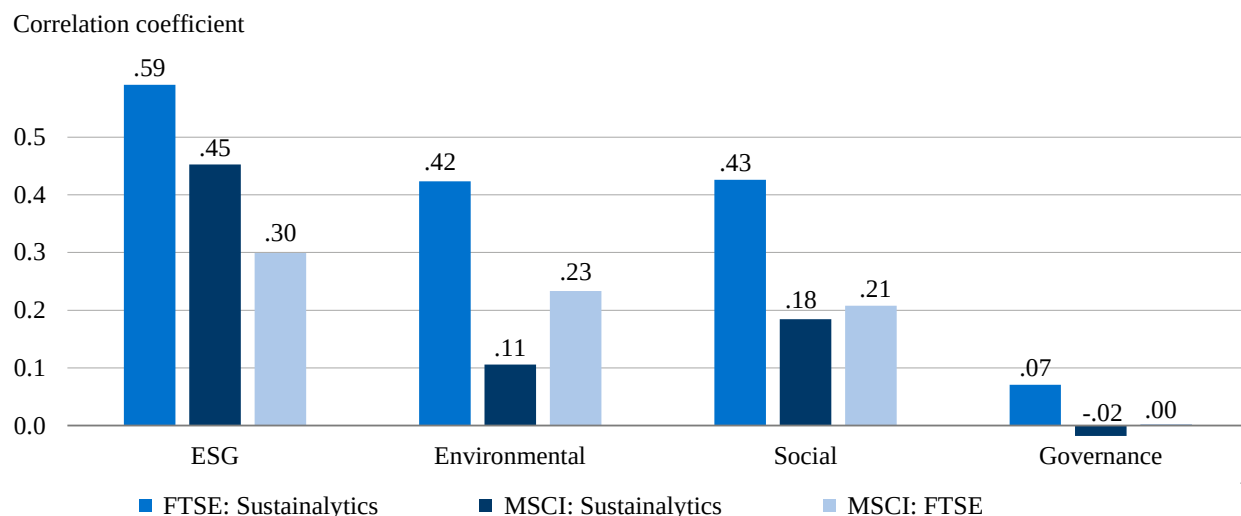
Source: Data from MSCI, FTSE Russell and Sustainalytics; analysis by the authors.

To explain the charts in Exhibit 2, consider the first company, Facebook. On the environmental pillar, Sustainalytics awards a very low ranking placing Facebook in the 1st percentile, while MSCI ranks it highly on the 96th percentile; on the social pillar the rankings are reversed, and MSCI places Facebook in the 7th percentile, while Sustainalytics places it in the 78th percentile. On the governance pillar, MSCI ranks three companies (JP Morgan Chase, Wells Fargo and Pfizer) as being extremely poor (percentiles between 4th and 7th) while Sustainalytics rates them as very high (percentiles between 87th and 99th). Even the overall ESG ratings can be strikingly different: MSCI rates Wells Fargo as poor (in the 12th percentile), while FTSE rates the company highly (in the 94th percentile).

Exhibit 3 shows the pairwise correlations between the three different raters, for their environmental, social and governance scores, and for their overall ESG ratings. These correlations are extremely low. In general, there is greater agreement between Sustainalytics and FTSE than between the other pairs of raters. Perhaps the most striking finding is the extraordinarily low correlations for governance. Raters disagree even on factual issues, such as whether the Chairman and CEO are separate, for which Berg, Koelbel, and Rigobon (2019) report a correlation of only 0.56.

As one would expect, the aggregate ESG ratings show more agreement than is demonstrated on the component parts. However, the average of the pairwise correlations is still a very low figure of just 0.45. Our findings are confirmed by other researchers including Berg Koelbel, and Rigobon (2019) and LaBella, Sullivan, Russell, and Novikov (2019). The ESG Data Guide reviews 117 services, and the consensus is that the correlations are generally low; see Environmental Finance (2020).

Exhibit 3: Correlations between ratings



Source: MSCI, FTSE Russell and Sustainalytics; analysis by the authors.

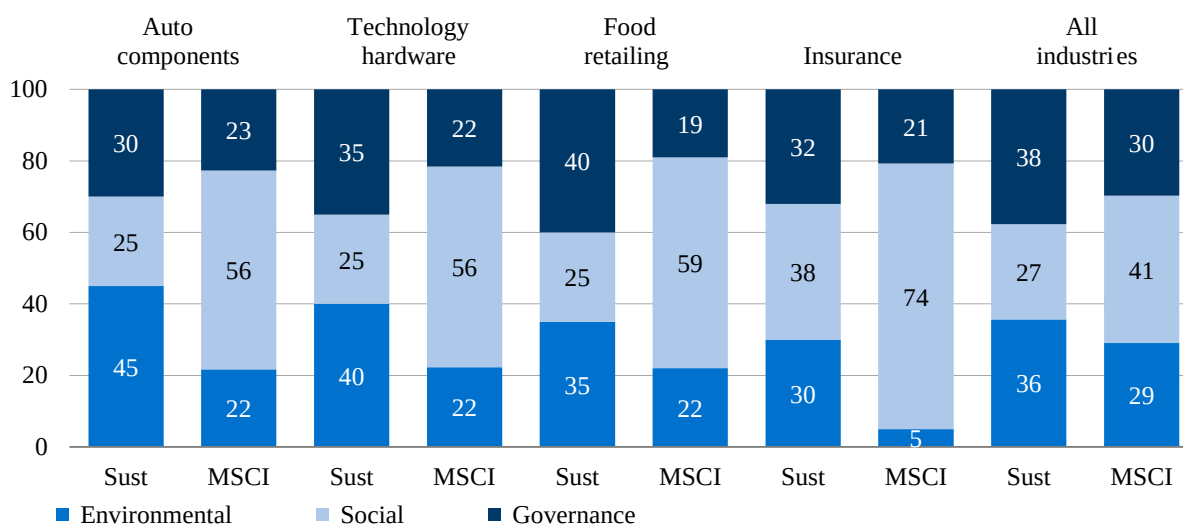
Explaining inconsistencies

In a recent survey, Kotsantonis and Serafeim (2019) highlight four factors that give rise to inconsistencies across ESG rating services. They are (1) data discrepancies, (2) benchmark choice, (3) data imputation, and (4) information overload. Considering them in turn, first, there is the variety and inconsistency of the metrics that purport to measure much the same thing. The diversity of measures gives rise to considerable dissimilarity in ratings reflecting firm-specific attributes, differing terminologies, metrics and units of measurement. Second, there are differences in how raters define the benchmark for comparisons. For example, Sustainalytics compares companies to constituents of a broad market index, whereas S&P compares companies to industry peers.

Third, at the company level, ESG ratings are plagued by missing data. When a company does not reveal metrics, some ESG raters assume the worst and (rather harshly) assign a score of zero. Others impute (somewhat generously) a score that reflects peers that do report the data. More sophisticated approaches use statistical models to estimate missing metrics, but are often unclear about why a company gets a low or high rating. Fourth, reflecting the never-ending expansion in the volume of public information and the lack of consensus on metrics, there is greater scope for raters to disagree about the scores for particular companies. Christensen, Serafeim, and Sikochi (2019) provide additional insights on rater disagreement.

An additional factor, not discussed by Kotsantonis and Serafeim, is the weighting scheme for ESG scores. As shown in **Exhibit 4**, Sustainalytics tends to place roughly equal weight on each of the three ESG pillars, whereas MSCI has a greater variability in the factor weightings that can give rise to material differences at the industry level. For the insurance industry, for example, Sustainalytics places roughly equal weightings on environmental, social and governance. In contrast, MSCI places a 5% weighting on environmental, 74% on social, and 21% on governance.

Exhibit 4: Industry weights by rater



Sources: Data from MSCI and Sustainalytics; analysis by the authors.

DOES SCREENING PAY OFF?

An understanding of ESG rankings and how and why they differ is important given their increasing usage. Indeed, there is a school of thought that positively slanting a portfolio toward responsible companies – positive screening – may be rewarded by better investment performance. Rankings also have a central role in ESG integration, which involves the systematic, explicit inclusion of ESG factors into investment analysis and portfolio selection.

Exhibit 5: Claims for an ESG premium, 2019

WSJ: “Rather than a mere window-dressing exercise conducted for the benefit of conscientious investors, investing with the environment in mind is now seen as a way to gain an edge. Funds are pouring billions of dollars into technologies and industries they think will benefit from a transition to a clean-energy world.”

FT: “For companies, ESG integration can provide a competitive edge, especially if rivals take a box-ticking approach to implementation. For investors, it can help them to beat the market through a discerning investment strategy — selecting companies that implement ESG well and avoiding ones that either “greenwash”... or invest in inconsequential practices.

WSJ: “Climate change is the growth story of our lifetime. We're not ideologically driven ... We want to make money. Climate investing is essentially a bet on a major repricing in markets”

FT: “I strongly believe that we can make good returns for clients while investing in companies that are doing well on ... ESG criteria”

WSJ: “It's no longer about avoiding the bad. It's about positively affirming the good and knowing that in doing so your financial returns will improve.”

Sources: FT = Financial Times; WSJ = Wall Street Journal (various dates)

ESG screening

Can screening lead to higher risk-adjusted returns? Certainly such claims abound, especially in the assertions of asset managers and the studies they cite to demonstrate this, and which they share with the public via the media. In **Exhibit 5**, we show some of the claims that have appeared in the financial media. The risk is that committed believers in the importance of ESG can become advocates, selecting, accepting and citing only evidence that confirms their beliefs.

The large and expanding body of studies on ESG often shows conflicting results, depending on the time frame and methodology. Compelling evidence – either way – extending over a sufficiently long period is in short supply. The data are often poor and do not yet go back far enough. This makes unequivocal judgements hard to reach.

When looking at equity returns, we must examine long periods of history as stock returns are very volatile and, even over multiple decades, we can still experience unusual returns. Dimson, Marsh, and Staunton (2020) show that it is hard to reach definitive conclusions on negative screening even using very long series.

We look first at corporate governance (G), then at environmental and social (E&S)

issues, and finally at climate change (a key part of E). In each case, we ask whether consideration of ESG data in investment approaches can pay off. Finally, we look at the performance of ESG funds and ESG indexes as a direct way of investigating whether there is a cost or benefit to ESG integration.

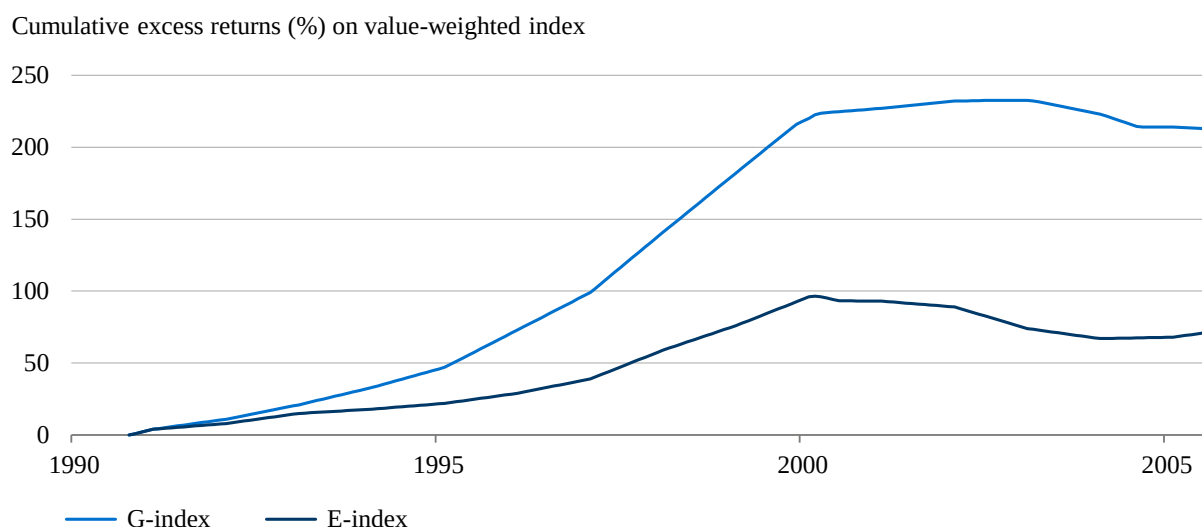
Corporate governance

Perhaps the most promising area for assessing whether ESG pays is the “G” of ESG. Corporate governance considerations have long been a major part of fund management. Larcker and Tayan (2019) suggest that, “Good governance is a set of processes or organizational features that, on average, improve decision-making and reduce the likelihood of poor outcomes.”

By this definition, good governance should lead to better corporate outcomes. Two caveats are in order. First, the measures used to assess good governance need to map closely onto Larcker and Tayan’s definition. It is not clear that ESG ratings achieve this. Second, investors may understand the opportunity, and compete away the profits from investing in well-governed companies.

Among the most influential and widely cited research in corporate governance is the U.S. study by Gompers, Ishii and Metrick (2003). This identified a corporate governance-based strategy centered on a G-index of 24 governance provisions that weakened shareholder rights. It involved buying the firms with the strongest shareholder rights and shorting the firms with the weakest. This provided abnormal returns of 8.5% per year during 1990–99. Bebchuk, Cohen, and Ferrell (2009) subsequently showed that these results were driven by just six of the 24 provisions and constructed a 6-provision E-Index.

Exhibit 6: The disappearing correlation with governance



Source: Bebchuk, Cohen and Wang (2013)

These findings generated considerable interest. They encouraged investors to emphasize good governance when selecting stocks, and they led to the development of many governance-based investment products. The attention paid to governance by the media, institutional investors, and researchers jumped sharply at the start of the 2000s and has remained high ever since. This was partly stimulated by the Gompers et al. paper, but was above all driven by a series of corporate governance scandals, including Enron, WorldCom and Tyco.

Learning effects

To the surprise of many people, the correlation between the G- and E-index governance indexes and abnormal returns disappeared during the 2000s. A paper by Bebchuk, Cohen and Wang (2013), which used the same methodology as the Gompers et al. study, found no association between the G- or E-indexes and abnormal returns during the 2000s or any sub-period within this decade. After outperforming the market during the 1990s, the governance-based strategy failed to outperform thereafter, as can be seen in **Exhibit 6**.

Bebchuk et al. attribute both the initial correlation and its disappearance to market participants learning to distinguish between good- and poor-governance firms, with respect to their levels of anti-takeover protection. Consistent with learning, the correlation's disappearance was associated with increased attention to governance by market participants.

Interestingly, Bebchuk et al. found that good governance firms continued, on average, to enjoy higher valuations and superior operating profits throughout both the decade of the 1990s and the 2000s. By the 2000s, however, this was fully priced in. The market had impounded relevant information into stock valuations, at least with respect to this particular measure of governance.

RESPONSIBLE INVESTING

Environmental and social (E&S) factors are at the heart of what is sometimes referred to as corporate social responsibility (CSR). Examples of the key issues include carbon and toxic emissions, effluent, use of clean energy, product carbon footprint, packaging material and waste, environmental stewardship, labor management and employee welfare, workplace safety, diversity and non-discrimination, human rights, product safety, sales practices, privacy protection and corporate philanthropy.

As we saw above, corporate governance, correctly measured, by definition impacts corporate performance. But the link between E&S and performance is less clear-cut. Friedman (1970) argued that “there is one and only one social responsibility of business – to use its resources and engage in activities designed to increase its profits so long as it stays within the rules of the game, which is to say, engages in open and free competition without deception or fraud.”

Friedman's concern was that managers would spend money on good deeds at the expense of focusing on their main responsibilities, and that they would do this on their own initiative, benefiting at the expense of shareholders. Some evidence that this is a reasonable fear, at least with corporate philanthropy, is provided in a paper by Cheng Hong, and Shue (2019). In Friedman's view, E&S spending is legitimate only if it adds value to the business. Well-run companies should recognize situations when investing in intangibles such as reputation, trust and good citizenship can add value – “doing well by doing good.” But if expenditures fail the

test of “doing well” (adding value), companies should leave “doing good” to others such as individuals, governments or non-governmental organizations, according to Friedman.

The business case for E&S expenditure is often strong. For example, the cost to aircraft manufacturers from issues in product safety, to automobile manufacturers from misleading emission tests, and to oil and chemical companies from failures in workplace safety can be high. Use of child labor in supply chains will certainly not endear companies to customers, who may reward instead companies that treat employees and suppliers well. Similarly, employees may work harder and contribute more if they are treated well and if they see that their employer cares about the community and the environment.

However, unlike corporate governance measures such as those mentioned above, social and environmental improvements can require large investments. For strip miners, tar sand extractors or coal-fired power station operators, the cost of a “clean up” could mean “doing worse by doing good” for a prolonged time.

E&S and financial performance

In a meta-analysis of 251 studies, Margolis, Elfenbein and Walsh (2009) find a positive but small effect of CSR on corporate financial performance. The average correlation was 0.13, but this fell to 0.09 for studies over the (then) most recent decade. There have been several other meta-studies, and even a meta-study of meta-studies by Friede, Busch and Bassen (2015), which summarized 2,200 individual ESG studies. More recent surveys include Busch and Lewandowski (2018), Eccles Kastrapeli, and Potter (2018), Giese and Lee (2019) and others.

In contrast to, say, pharmaceutical studies, these meta-studies give the same weight to poor quality, unpublished papers as to high-quality refereed articles in top journals (the latter being in a small minority). They lump together accounting and market-based performance measures, although the latter generally reveal smaller effects. Where market-based measures are used, they are seldom appropriately adjusted. Two-thirds of studies suffer from look-ahead bias by comparing coincident measures of CSR and financial performance.

At best, the studies covered indicate a correlation between CSR and corporate financial performance, with no indication of causality. We cannot say whether firms that do good do well, or whether firms that do well do good. Nor can we say whether the correlation comes from some unidentified variable that impacts both ratings and financials. These studies are also weakened by the assumption that E&S scores adequately measure responsible behavior.

Does CSR pay?

Rather than looking at correlations, Krüger (2015) identifies 2,116 CSR events from 2001–07 for 745 U.S. companies from the KLD (now MSCI) database. He examines abnormal performance around the day the events become public. Negative events, especially those related to the environment or community, have a damaging impact on stock value. This suggests a substantial cost to corporate social irresponsibility. For positive events, investors react slightly negatively, but the reaction is much weaker. Analysis of the positive events reveals that investors respond negatively when there are greater agency problems, and positively when the firm is seeking to remedy an earlier unfavorable event. Events with stronger legal or economic implications generate larger abnormal returns.

Other studies find that CSR information is impounded into stock prices with increasing speed. Halbritter and Dorfleitner (2015), using three rating providers and a sample period from

1990 to 2012, found that for all three raters the outperformance of a highly-rated over a lowly rated portfolio was large and positive from 1990 to 2001, about half the size from 2002 to 2006, and completely absent from 2007 to 2012. This pattern resembles the findings for corporate governance. Investors learn over time, and markets become more efficient at discounting E&S information. The implications of ESG ratings may now be fully reflected in stock prices.

CLIMATE CHANGE

Climate change is the key environmental issue within ESG. There is consensus that this is a premier global challenge, perhaps the greatest mankind faces. It is now a central concern of investors, and climate-risk analysis is now part of the mainstream of asset management. The threat from climate change is clear, and this provides investors with an opening to exert influence for good. It may also offer investment opportunities. Two possibilities that are regularly suggested are (1) shunning (or shorting) likely “stranded assets” and (2) focusing on low-, rather than high-carbon investments.

Stranded assets

The 2015 Paris Agreement set out an international framework to limit global warming to below 2°C and ideally below 1.5°C. If this target is to be achieved, most of the world’s existing fossil fuel reserves will have to remain in the ground. According to McGlade and Ekins (2015), about a third of worldwide oil reserves, a half of gas reserves, and over three-quarters of coal reserves would remain unused, as abandoned “stranded assets.”

Not only would reserves be left behind, but resources – past and future – committed to extraction and processing would be unneeded. Assets would be “stranded” to the extent that regulatory change and legal pressures inhibit future consumption of traditional fuels while the cost of renewable energy is expected to fall. It has been suggested that the risk of assets turning out to be stranded could be large, and that financial markets fail to recognize this.

If investors do fail to recognize this, stranded assets may indeed impose losses. However, this would also provide an opportunity for better informed investors to sell, or short, such holdings. In competitive financial markets investors are about as likely to underestimate as to overestimate asset values – in other words, risk is roughly symmetric. Furthermore, investors who were unaware of climate risk early in the 21st century have now had plentiful opportunities to impound this risk into current market prices. It follows that, as well as downside risk, stranded assets may provide upside potential to investors.

Consequently, avoidance of potentially stranded assets could be financially costly. This has happened in the past. Tobacco stocks faced the prospect of litigation and changes in customer tastes, but divestment cost CalPERS and Norway’s GPF \$0.65 billion and \$1.9 billion, respectively, over a six-year period; see Barber (2007) and Kasikci and Rajasingam (2017). Assuming they are able to identify potentially stranded assets, there is no evidence that climate experts have superior ability to appraise stock market values.

Low-carbon

The attention of many investors is on climate change, and the existential threats posed by environmental degradation; see Chambers, Dimson, and Quigley (2020). Asset owners and

investment professionals want access to information on the carbon exposure of their investee companies. This is essential if they are to exert influence.

It is much less clear, however, whether this information is linked to investment performance. Sadly, the simple question “is there a green factor premium (a reward for sustainable investing) or a carbon factor premium (a reward for tolerating environmental damage)?” has defied resolution. Recent academic studies disagree on the effects that a firm’s carbon emissions have on its stock performance, despite using the same data on emissions, from Trucost starting in 2005.

These studies suffer from methodological issues. Moreover, they are restricted to the short time interval dictated by emissions data availability: see Atta-Darkua, Chambers, Dimson, Ran, and Yu (2020). In a careful appraisal of conflicting empirical evidence, Lioui (2019) argues that we are still in the dark as to the existence of any carbon anomaly or pricing factor and that “we are very far from having convincing evidence.”

ESG INVESTMENT PERFORMANCE

Nitsche and Schröder (2018) examined 186 European and global ESG funds, matching their holdings to three ESG rating services. They found that the funds were not “closet” conventional funds, but, on average, exhibited a significantly higher ESG score. However, there was almost a one-in-three chance that a randomly selected ESG fund would have a lower ESG rating than a randomly selected conventional fund.

It is clear that a comparison of ESG versus non-ESG investing requires careful consideration of the empirical evidence. A direct way to investigate whether there is a cost or benefit to ESG investing is to study the performance of ESG funds. We consider this approach first, and later turn to the returns from ESG indexes.

Fund performance

Studies of sustainable fund performance date back to the 1970s and at the time mostly focused on exclusions. Earlier work typically concluded that there was little performance difference between ESG and conventional funds. However, they mostly looked at small samples, short time periods, specific countries and performance measures that failed to adjust for factor exposure. One of the first comprehensive studies, by Renneboog, Ter Horst, and Zhang (2008), examined 440 ESG funds in 17 countries, comparing them with 16,000 conventional funds. The ESG funds in the U.K., USA and in many European and Asia-Pacific countries underperformed their domestic benchmarks by 2.2%–6.5%. Compared with conventional funds in most countries, however, the underperformance was not statistically significant. Subsequent research has yielded similar findings.

While most studies simply compare ESG with conventional funds, El Ghouli and Karoui (2017, 2019) adopted a novel approach. They selected a sample of 2,168 U.S. equity funds without regard to whether they claimed to be ESG funds or not. They then constructed an asset-weighted composite CSR score for each fund and found that higher scoring funds had poorer and more persistently poor performance than their lower-scoring counterparts. In their later paper, they examine how closely funds track an ESG index. Somewhat in contrast to their earlier work, they find that, for a sample of 2,516 U.S. mutual funds over the period 2010–2017, their “results are consistent with the hypothesis that SRI does not significantly damage fund performance.”

To sum up, despite the extensive literature with its varying findings, we find almost no convincing studies showing that ESG funds outperform on a sustained basis. Most studies find neutral to mildly negative relative performance by ESG funds. In **Exhibit 5** above, we cited claims that ESG investing leads to superior performance and/or lower risk. The ESG fund performance literature would struggle to support this. However, the balance of evidence suggests that the price that ESG investors pay for their principles is probably quite modest.

Index performance

As an alternative to researching ESG funds, several studies have examined the performance of ESG indexes. This is attractive as the indexes are used as benchmarks by ESG investors, they apply the same types of screening (both negative and positive) as most funds, and performance is not distorted by costs, fees, timing or stock selection. However, ESG indexes vary in their construction rules and in how they deal with regional coverage and data biases connected to company size.

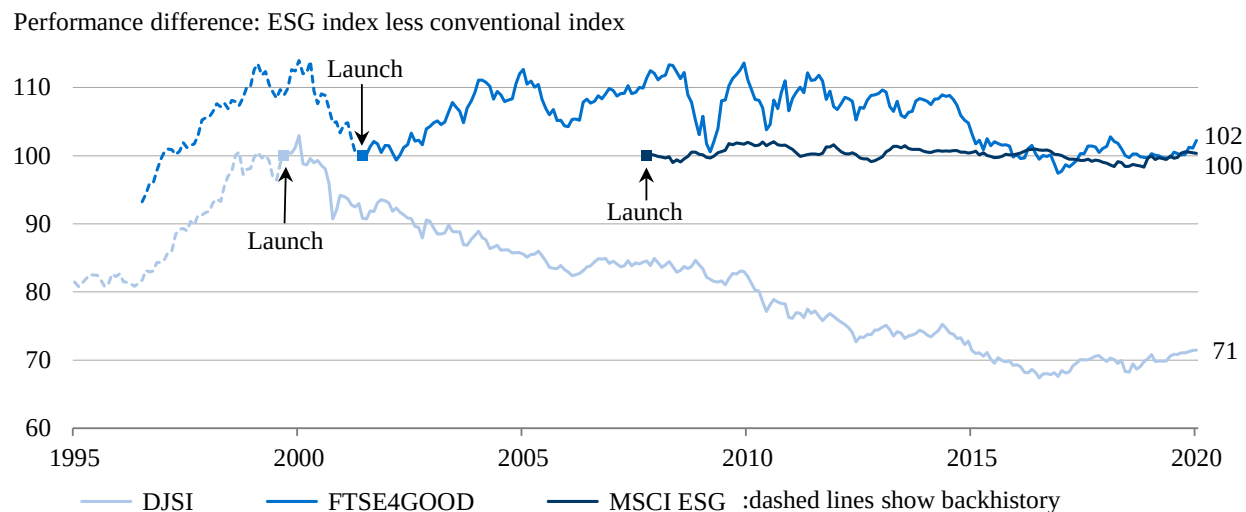
Schröder (2007) examined 29 ESG equity indexes from 11 suppliers and compared their performance with the closest available conventional indexes. He found neutral relative performance, although most of the indexes had a higher risk than their benchmarks. He also found clear evidence of look-back bias: in every case where the index history incorporated a pre-launch back-history, performance over the back-history exceeded that over the post-launch period.

To provide an update to Schröder's classic work, we examine the flagship global ESG indexes of three major index providers: the MSCI World ESG Leaders, FTSE Russell's FTSE4GOOD and S&P's Dow Jones Sustainability World index (DJSI). Using similar approaches, MSCI and FTSE Russell both exclude sin stocks and then select on ESG scores to target the highest-rated 50% free-float market capitalization of each sector of the parent index (the MSCI World and the FTSE Developed Markets index). DJSI adopts a rules-based selection of the top 10% by number of the most sustainable companies in the S&P Global BMI index, based on ratings from SAM (now owned by S&P).

While recognizing the differing index construction methodologies and related impacts, we compare each index to its conventional counterpart and plot the difference in cumulative performance. We do this from the launch date to end-2019. The plot is shown in **Exhibit 7**. After going live, both the FTSE4GOOD and MSCI ESG indexes experienced neutral relative performance (see the solid line plots). In contrast, since its August 1999 launch, the DJSI has underperformed the S&P Global BMI by 29%, equivalent to 1.6% per year.

Both the DJSI and the FTSE4GOOD have pre-launch back-histories (see dashed lines). Consistent with Schröder's finding, these indexes outperformed over the period from the start of the back-history until the launch date. Schröder claims that the back-histories for the indexes he examined were usually calculated retrospectively using the index composition at the launch date, which would virtually guarantee spurious outperformance because of survivorship and success bias. An alternative explanation is that index compilers use pre-launch data to develop a methodology with an attractive (albeit artefactual) performance record.

Exhibit 7: Performance of Global ESG indexes (in USD)



Source: Data from MSCI, FTSE Russell, and Dow Jones via Refinitiv.

We find no evidence of ESG outperformance. Equally, however, for the two indexes that most closely approximate to what ESG investors actually do (namely, FTSE4GOOD and MSCI-ESG) there was no evidence of underperformance either.

CONCLUSION

This article has focused on ESG ratings. There is now a vast demand for sustainable and socially aware investment products and services. It would be difficult for any asset manager to satisfy this demand unless data is sourced, at least in part, from external service providers, and there is now a wide array of ESG rating agencies. We have described their advantages and weaknesses. In particular, there is little relation between the ESG appraisals of a given stock by alternative agencies.

Clearly, therefore, ESG ratings should not be treated as a black box or used mechanically. Blanket use of ESG scores is not the solution. At best, they are a starting point. Analysts and fund managers need to understand how they are constructed and supplement them with their own scrutiny to build a holistic understanding of a company and hence ensure their investments best reflect the values of their end-investors.

It is often claimed that the use of ESG ratings can enhance risk-adjusted performance. For investors, there may well be, or at least have been, financially worthwhile opportunities from investing in companies based on their ESG credentials at least for a period of time when prescient investors have identified price relevant characteristics that have not been fully recognized by the market. But such benefits are likely to be short-lived thanks to competition from other investors and by peers catching up. We have seen evidence in both the corporate governance and E&S fields that investors learn, and that this learning is in due course reflected in valuations, making markets more efficient.

For a long-term investor, the perspective we take is that we lack unambiguous evidence that ESG screening enhances expected return or reduces risk. This remains true whether we look at the performance of companies based on their ratings, or at ESG funds or indexes. Equally, however, we find no strong evidence of underperformance. For ESG investment strategies based on exclusions, theory and evidence available so far suggest that the sacrifice made by ESG investors is a slight shortfall in expected return and a minor reduction in diversification. The price for ethical principles appears small, and one that many virtuous investors may be content to bear.

ACKNOWLEDGEMENTS

This article draws on research presented in Dimson, Marsh, and Staunton (2020). The authors thank Nannette Hechler-Fayd'herbe, Richard Kersley, and colleagues at Cambridge University and London Business School for helpful comments and shared insights.

REFERENCES

- Atta-Darkua, V., Chambers, D., Dimson, E., Ran, Z., and Yu, T., 2020. Strategies for responsible investing: Emerging academic evidence. *Journal of Portfolio Management: Ethical Investing* 46(3): 26–35.
- Barber, B., 2007. Monitoring the monitor: Evaluating CalPERS' activism. *Journal of Investing* 16(4): 66–80.
- Bebchuk, L., Cohen, A., and Ferrell, A., 2009. What matters in corporate governance? *Review of Financial Studies*, 22: 783–827.
- Bebchuk, L., Cohen, A., and Wang, C., 2013. Learning and the disappearing association between governance and returns, *Journal of Financial Economics*. 108(2): 323–348.
- Berg, F., Koelbel, J., and Rigobon, R., 2019. Aggregate confusion: The divergence of ESG ratings. *MIT Sloan Research Paper* No. 5822-19. SSRN.com/id=3438533
- Busch, T., and Lewandowski, S., 2018. Corporate carbon and financial performance: A meta-analysis, *Journal of Industrial Ecology* 22(4): 745–759.
- Chambers, D., Dimson, E., and Quigley, E., 2020. To divest or to engage: A case study of investor responses to climate activism. *Journal of Investing: ESG Special Issue*. 29(2): 10–20.
- Cheng, I., Hong, H., and Shue, K., 2019. Do managers do good with other peoples' money? Fama-Miller Working Paper No. 12–47. SSRN.com/id=1962120.
- Christensen, D., Serafeim, G., and Sikochi, S., 2019, Why is corporate virtue in the eye of the beholder? The case of ESG ratings, Harvard Business School Working Paper.
- Dimson, E., Marsh, P., and Staunton, M., 2020. ESG investing. In: *Credit Suisse Global Investment Returns Yearbook 2020*. Zurich: Credit Suisse Research Institute. 47–64.
- Eccles, R., Kastrapeli, M., and Potter, S., 2018. How to Integrate ESG into investment decision-making: Results of a global survey of institutional investors, *Journal of Applied Corporate Finance* 29(4): 1–137.
- El Ghoul, S., and Karoui, A., 2017. Does corporate social responsibility affect mutual fund performance and flows? *Journal of Banking and Finance* 77: 53–63.

- El Ghoul, S., and Karoui, A., 2019. Fund performance and social responsibility: New evidence using social active share and social tracking error. University of Alberta working paper. SSRN.com/id=3489201.
- Environmental Finance, 2020. *Guide to ESG data providers*. Field Gibson Media. www.environmental-finance.com/content/guides/guide-index
- Friede, G., Busch, T., and Bassen, A., 2015. ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment* 5(4): 210–233.
- Friedman, M., 1970. The social responsibility of business is to increase its profits. *The New York Times Magazine*. (18 September)
- Giese, G., and Lee, L., 2019. Weighing the evidence: ESG and equity returns. MSCI ESG Research. (April).
- Gompers, P., Ishii, J., and Metrick, A., 2003, Corporate governance and equity prices. *Quarterly Journal of Economics*. 118(1): 107–155.
- Halbritter, G., and Dorfleitner, G., 2015. The wages of social responsibility – Where are they? A critical review of ESG investing. *Review of Financial Economics* 26 (3): 25–35.
- Kasikci, F., and Rajasingam, S., 2017. The Norwegian Government Pension Fund Global: The cost of ethical exclusions. Oslo: BI Norwegian Business School.
- Kotsantonis, S., and Serafeim, G., 2019. Four things no-one will tell you about ESG data. *Journal of Applied Corporate Finance* 31(2): 50–58.
- Krüger, P., 2015. Corporate goodness and shareholder wealth. *Journal of Financial Economics*. 115(2): 304–329.
- Larcker, D., and Tayan, B., 2019. Loosey-goosey governance: Four misunderstood terms in corporate governance. *Stanford Closer Look Series*, No. CGRP-79. SSRNcom/id=3463958.
- LaBella, M., Sullivan, L., Russell, J., and Novikov, D., 2019. *The Devil Is in the Details: Divergence in ESG Data and Implications for Sustainable Investing*. QS Investors.
- Lioui, A., 2019. The decarbonisation factor: A new academic fiction? A tale of two papers. *EDHEC-Scientific Beta Working Paper* (December).
- Margolis, J., Elfenbein, H., and Walsh, J., 2009. Does it pay to be good... and does it matter? A meta-analysis of the relationship between corporate social and financial performance. Working paper, Harvard Business School. SSRN.com/id=1866371.
- McGlade, C., and Ekins, P., 2015. The geographical distribution of fossil fuels unused when limiting global warming to 2° C. *Nature* 517, 187–190 (2015).
- Nauman, B., 2019. S&P acquires ESG ratings arm of RobecoSAM, *Financial Times* (21 November).
- Nitsche, C., and Schröder, M., 2018. Are SRI funds conventional funds in disguise or Do they live up to their name? In Boubaker, S., Cumming, D., and Nguyen, D. [eds.], *Research Handbook of Investing in the Triple Bottom Line: Finance, Society and the Environment*. Edward Elgar Publishing Inc. ch19: 414–446.

Renneboog, L., Ter Horst, J., and Zhang, C., 2008. The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance* 14(3): 302–22.

Schröder, M., 2007. Is there a difference? The performance characteristics of SRI equity indexes. *Journal of Business Finance & Accounting*. 34 (1–2): 1–401.

Walter, I., 2019 Sense and nonsense in ESG scoring. *Stern School working paper*, NYU SSRN.com/id=3480878.