Understanding the Business Relevance of ESG Issues

George Serafeim and Aaron Yoon

In the last decade, an increasing number of companies have released environmental, social and governance (ESG) data, fueled both by voluntary disclosure and regulatory mandates. This came, partly, as a response to increasing demand by investors to understand how companies were managing issues that could be potentially material for business performance. In line with that, a growing literature suggests that ESG issues are economically meaningful thereby associated with stock returns (Edmans 2011; Dimson et al. 2015; Lins, Servaes and Tamayo 2017; Bolton and Kacperczyk 2020; Dai et al. 2020; Welch and Yoon 2021).

However, because of the multidimensionality of the ESG construct, which includes a wide variety of issues and data ranging from water withdrawal and carbon emissions to lost time injury rates and percentage of women in management, a prioritization of the different issues and metrics was needed. From a management's perspective, this is prudent given that managerial attention is a scarce resource. From an investor's perspective, this can be helpful given that it is highly unlikely that all ESG issues and data are investment relevant. Against this backdrop, a great level of interest was developed around the idea of materiality, meaning which ESG issues are likely to be more or less important across different companies.

What is Business Relevance?

In the process of prioritizing among different ESG efforts, organizations have been releasing materiality matrices for years (see Figure 1 for example). On the one axis you have importance to society and on the other axis you have importance to the company. This matrix nicely depicts the issue of double materiality

that has become increasingly important in regulatory discussions, as embraced by the European Union.[1] Inside the matrix, scattered across the plot, are different ESG issues. As one might expect, firms in different industries have vastly different ESG issues that score low or high on how important they are for the company or society.

For example, for Cisco, data security and privacy issues would be very important to the company as well as to the stakeholders (see Figure 1 Panel A). For Unilever, packaging and waste issues would be very important to the company as well as to the stakeholders. As such, we can expect community relation around mining sites to be a very important issue for mining firms while climate transition risk and emission issues to be important for energy companies. As unlikely it is that every ESG issue is important for the performance of most companies, it is also very unlikely that no issues are important for most companies.

Identifying Business Relevance

Most companies have been identifying relevance to company performance through stakeholder engagement, mapping the different issues according to their business strategy, unique competitive positioning, and resources (Eccles, Ioannou and Serafeim 2014). While those approaches can inform individual firm-level action they are unlikely to be helpful in providing comparable information for the assessment of company performance. In response, the Sustainability Accounting Standards Board (SASB) developed standards that identify the most important issues at the industry level (from the perspective of financial materiality - meaning importance to investors), providing reporting guidance for companies based on industry membership.

Standards are needed because, among other benefits, they can increase the comparability of the reported information, something that has been a remarkable weakness in the ESG reported information (Amel-Zadeh

[1] Double materiality refers to the idea of considering not only the materiality of an ESG outcome on the company (financial materiality) but also the materiality of a company's actions on an ESG outcome (societal materiality). For the rest of the paper when we refer to materiality, we refer to financial rather than societal materiality.

and Serafeim 2018). However, it is important to note that standards impose a degree of homogeneity at the level of analysis that the standard is applied (i.e., industry) thereby masking significant heterogeneity across firms in how ESG issues might be more or less important based on geography, value chain choices, business model differences and customer profiles. This trade-off between comparability and accuracy is well understood by accounting scholars (e.g., Demski 1974; Dopuch and Pincus 1984; Dye and Verrecchia 1995; Fang, Iselin, and Zhang 2017).

Evidence of Relevance

Assessing the differential importance of different ESG issues, investment managers develop their own proprietary classifications and models. In contrast, researchers have relied on publicly available standards, such as SASB, and they have found that standards can be helpful in assessing the differential importance of ESG issues. Using different data sources, time periods, and samples, various studies have shown that some ESG issues are more important than others for assessing business performance.

For example, Khan, Serafeim and Yoon (2016) find, using KLD data, that firms improving their scores on ESG issues that are classified as financially material by SASB perform better than firms with decreasing scores on material ESG issues. This relation is not observed for data that are not classified as financially material. Grewal, Hauptmann and Serafeim (2020) show that material ESG disclosure scores constructed using Bloomberg data but also data from SASB's disclosure database are associated with lower stock price synchronicity. However, they do not find such evidence for immaterial ESG disclosure scores. Moreover, they document that following the publication of materiality standards by SASB, firms that increase their disclosure on material ESG issues experience a decline in stock price synchronicity.

Flammer (2020) shows positive abnormal stock reactions at the time of issuance of green bonds for firms in industries where environmental issues are classified as a financially material issue by SASB. Chen, Dong, and Lin (2020) show that increases in institutional shareholder holdings are associated with improvements

in sustainability scores especially on material ones suggesting that investors recognize the importance of materiality. Grewal, Riedl and Serafeim (2018) show that around the announcement of a nonfinancial disclosure directive firms with poor sustainability disclosure had negative abnormal returns and that this negative reaction was even stronger for firms that sustainability issues are more material, as proxied by MSCI sustainability risk exposure assessments. Henisz and McGlinch (2019) show a relationship between material sustainability scores, where materiality is specified by SASB, and credit risk, including credit yield spread and credit default spread.

Serafeim and Yoon (2022) show abnormal stock market reactions around the announcement of ESG events classified as material by SASB and no reactions otherwise. Moreover, they show that these reactions are even greater the more salient those events are. McGlinch and Henisz (2020) show that firms which are better able to manage ESG issues financially material to their industry, as identified by SASB, are more likely to report unexpected increases in revenue, favorable reductions in operating costs, fewer legal or regulatory interventions, gains in employee productivity, and interest in increasing both organic and inorganic business investment. Matsumara, Prakash, and Vera-Munoz (2020) use SASB's materiality classification to infer market expectation about climate risk materiality and find that firms that disclose climate risk issues exhibit lower cost of capital. Spandel, Schiemann and Hoepner (2020) find a negative (positive) capital market reaction for firms with low (high) material ESG performance following SASB's publication of standards.

A Measurement Errors Framework

In many ways, observing a relation between ESG and firm performance can be surprising. This is because most ESG issues are unlikely to be financially material for a company and thereby creating aggregate scores across a wide variety of them is likely to yield statistics with very low signal to noise ratio from a business performance perspective. Relying on standardized classifications, such as SASB, to isolate likely material

issues is helpful but it is also likely to embed large measurement errors. Here is a simple way of understanding that phenomenon:

$$\text{True\_Materiality}_{ikjt} = \text{Standards\_Materiality}_{kj} + \kappa_{ij} + \mu_{ijt} + \lambda_{ik} + \nu_{kjt}$$

The true materiality of a specific sustainability information j for firm i in industry k at time t is unobservable. Through standards such as SASB, what a researcher can observe is a materiality classification for industry k and sustainability information j. Then it follows that there are several error terms in the relation between the true and observed materiality. We outline them below:

K_{ij} denotes the fact that the sustainability data item is just a proxy for the information identified as material by SASB, as most SASB data items have not been reported historically by companies and even the data items identified by SASB might not necessarily be the most valuable indicator for assessing business performance. Most researchers utilize some type of ESG scoring mechanisms that are provided by different commercial data firms. A well-known fact now is the low agreement between different providers in their assessments of similar ESG issues (Christensen, Serafeim and Sikochi 2021; Berg, Kolbel, and Rigobon 2021) and that many ESG evaluations have inherent limitations due to missing information, data imputations, and reference benchmarks (Kotsantonis and Serafeim 2018).

μ_{ijt} denotes the fact that the sustainability data item j might not be material for firm i even though that firm is in industry k because the firm's operations or strategy might differ. For example, materiality might differ across subindustries within the same industry or even across companies within a subindustry. Even though labor standards in the supply chain might be an important issue for apparel companies, one company that has most of its supply chain in high human rights risk areas, such as Bangladesh, is likely to have a very different exposure compared to an apparel company with a supply chain, in a low-risk area, such as Canada.

λ_{ik} denotes the fact that a firm might be misclassified from an industry perspective as it relates to material sustainability issues and therefore its materiality classifications might be misjudged. Here the issue becomes

even more important given that many companies are industrially diversified and therefore inherently might be adhering more or less to the industry standard.

v_{jkt} denotes the fact that materiality can be dynamic and the materiality for sustainability information j in industry k varies over time. For example, think about diversity and inclusion following the MeToo movement or Black Lives Matter, which have likely increased the materiality of these issues. Similarly, changes in social sentiment towards firms that emit large amounts of carbon are likely changing the financial materiality of carbon emissions.

The point here is that there are many error terms that could be dampening the relation between the observed materiality classifications and the true materiality of a specific ESG data item. As a result, inability to document statistically significant results could be the result of one or more of those errors in measurement that can bias coefficients towards zero. Consistent with this, recent research shows that using instrumental variables to reduce errors-in-variables significantly increases the association between ESG data and stock returns (Berg, Kolbel, Pavlova and Rigobon 2021).

Models estimated in Berchicci and King (2021) suggest that making different variable construction, sample period, and control variable choices can yield different results with regards to the relation between ESG scores and business performance. Given the discussion above, this is certainly highly plausible, especially when one utilizes crude ESG measures, as in KLD's binary classification of strengths and weaknesses. However, not all models are created equal and failing to understand the implications of different model choices can lead predictably to wide variation in estimated coefficients.

For example, Khan, Serafeim and Yoon (2016) use a dichotomous instead of a continuous measure because of the weaknesses of ESG data and the crudeness of the KLD data, which is a series of binary variables. Creating a dichotomous variable (i.e., top quintile for example) could be well suited when trying to identify firms on a specific characteristic and the metric identifying that characteristic is likely to be noisy. A continuous measure assumes that for the whole sample researchers can be confident in the distance that

each firm exhibits from each other. Therefore, the use of continuous measure is likely to lead to significantly weaker results, as in Berchicci and King (2021), because it exacerbates κ_{ij}. Similarly, the different materiality mappings between SASB and KLD can lead to larger measurement error due to κ_{ij} if researchers allocate KLD data items that are immaterial to the material group. This is especially true given that there are only a few issues within industry and as a result data items that are material, therefore even small misallocations could change ESG assessments significantly. Conducting sub-period analyses, that utilize KLD data when for that whole period, post 2010, the KLD product had been discontinued and data collection efforts for KLD had ceased (Eccles, Lee and Stroehle 2020) is likely to yield, unsurprisingly, noisy estimates given the deteriorating quality of the data. In addition, the different industry classifications utilized in Berchicci and King (2021) can lead to larger measurement error in λ_{ik} further lowering the power of the test. Researchers should carefully make choices, reflecting on the different types of measurement errors that those choices could exacerbate or mitigate.

Implications for Future Research

In the last few years, the literature has developed a better understanding of the financial materiality of ESG issues. The first significant point is that financial materiality of ESG issues is likely to be dynamic, making static classifications (e.g., SASB's classification) likely weak tests of the underlying structure between ESG metrics and financial performance (Freiberg, Rogers and Serafeim 2020; Kuh, Shepley, Bala and Flowers 2020). In essence, the relationship between different ESG issues and financial performance is likely to vary over time. Taking that into account is likely to mitigate ν_{jkt} mentioned above.

The second point is that given materiality is a dynamic concept that evolves over time, ESG disclosure will be more difficult to regulate compared to financial disclosure (Freiberg, Rogers and Serafeim 2020). Policy makers might need a new more flexible, principles-based approach to regulating ESG disclosure and measurement as different practices and new dynamics emerge. Comparability of information might be more

challenging to achieve as the dynamic nature of materiality will manifest at different points in time and with different intensity across companies and industries.

The third point is that we have made progress in improving the quality of ESG data and the field is continuously improving, as companies are releasing higher quality data, data providers are improving quality controls, advancements in technology allow for the creation of new measures from big data, and regulatory structures are improving reporting practices. Therefore, researchers should use state-of-the art ESG data rather than rely on legacy datasets, such as KLD, that are already phased out. While prior research, has used datasets such as KLD, primarily because of their long time-series and in the absence of alternatives, this came at the cost of utilizing data with subjective binary evaluations if a firm exhibits a strength or a weakness on a specific ESG issue. Newer datasets provide more granular data and access to raw key performance indicators with now also considerable time series. For example, Factset's TruValue Labs dataset collects ESG news data from over 100,000 sources and use national language processing techniques to create ESG sentiment scores. Of course, such datasets could also come with significant challenges given the noise embedded in big data. However, to increase transparency and validate the data, it allows users to track the original source of the articles and events that inform the sentiment analysis for each specific issue (Serafeim and Yoon 2021). Newer and more robust datasets are likely to yield more insights. This would likely mitigate κ_{ij}.

The fourth point is that firm-specific evaluations are likely to yield much more precise assessments of materiality if companies differ significantly in their business model and strategic choices thereby yielding unique ESG risks and opportunities. This will be especially important as companies are trying to differentiate themselves through the adoption of unique practices (Ioannou and Serafeim 2019). In turn, such evaluations would reduce μ_{ijt}. This relates to further mitigating λ_{ik} which can arise from the disconnect between industrial memberships and their accuracy in reflecting the business exposures of a firm.

Conclusion

Analyzing the importance of ESG data for business performance is likely to continue to be a growing field of inquiry. In our opinion, the study of ESG data needs to be done with a deep appreciation of the complexities, nuances, and time evolution of the underlying construct. It is the collection of evidence across multiple studies using different data, samples, contexts, time periods along with understanding developments in practice that can lead to a robust understanding of important phenomena.

References

Amel-Zadeh, A. and Serafeim, G., 2018. Why and how investors use ESG information: Evidence from a global survey. Financial Analysts Journal, 74(3), pp.87-103.

Berchicci, L. and King, A., 2020. Material Sustainability Ratings and Shareholder Returns: A Model Uncertainty Analysis. Working Paper.

Berg, F., Koelbel, J.F. and Rigobon, R., 2020. Aggregate confusion: the divergence of ESG ratings. Available at SSRN 3438533.

Bolton, P. and Kacperczyk, M., 2020. Do investors care about carbon risk? (No. w26968). National Bureau of Economic Research.

Chen, T., Dong, H. and Lin, C., 2020. Institutional shareholders and corporate social responsibility. Journal of Financial Economics, 135(2), pp.483-504.

Christensen, D., Serafeim, G. and Sikochi, A., 2019. Why is Corporate Virtue in the Eye of The Beholder? The Case of ESG Ratings. Working Paper.

Dai, R., Liang, H. and Ng, L., 2020. Socially responsible corporate customers. Journal of Financial Economics.

Demski, J.S., 1974. Choice among financial reporting alternatives. The Accounting Review, 49(2), pp.221-232.

Dimson, E., Karakaş, O. and Li, X., 2015. Active ownership. The Review of Financial Studies, 28(12), pp.3225-3268.

Dopuch, N. and Pincus, M., 1988. Evidence on the choice of inventory accounting methods: LIFO versus FIFO. Journal of Accounting Research, pp.28-59.

Dye, R.A., Verrecchia, R.E., 1995. Discretion vs. uniformity: Choices among GAAP. The Accounting Review 70, 389-415.

Eccles, R.G., Lee, L.E. and Stroehle, J.C., 2020. The social origins of ESG: An analysis of Innovest and KLD. Organization & Environment, 33(4), pp.575-596.

Edmans, A., 2011. Does the stock market fully value intangibles? Employee satisfaction and equity prices. Journal of Financial economics, 101(3), pp.621-640.

Fang, V.W., Iselin, M. and Zhang, G., 2017. *Financial statement comparability: benefits and costs* (No. 2858301). SSRN Working Paper.

Flammer, C., 2020. Corporate green bonds. Journal of Financial Economics,

Freiberg, Rogers and Serafeim 2020. Pathways to materiality: How sustainability issues become financially material to corporations and their investors. Working paper.

Grewal, J., Hauptmann, C. and Serafeim, G., 2020. Material sustainability information and stock price informativeness. Journal of Business Ethics, pp.1-32.

Grewal, J., Riedl, E.J. and Serafeim, G., 2019. Market reaction to mandatory nonfinancial disclosure. Management Science, 65(7), pp.3061-3084.

Henisz, W.J. and McGlinch, J., 2019. ESG, Material Credit Events, and Credit Risk. Journal of Applied Corporate Finance, 31(2), pp.105-117.

Kotsantonis, S. and Serafeim, G., 2019. Four things no one will tell you about ESG data. Journal of Applied Corporate Finance, 31(2), pp.50-58.

Khan, M., Serafeim, G. and Yoon, A., 2016. Corporate sustainability: First evidence on materiality. The accounting review, 91(6), pp.1697-1724.

Kuh, T., Shepley, A., Bala, G. and Flowers, M., 2020. Dynamic Materiality: Measuring What Matters. Available at SSRN.

Lins, K.V., Servaes, H. and Tamayo, A., 2017. Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. The Journal of Finance, 72(4), pp.1785-1824.

Matsumura, E.M., Prakash, R. and Vera-Munoz, S.C., 2020. Climate Risk Materiality and Firm Risk. Available at SSRN 2983977.

McGlinch, J. and Henisz, W.J., 2020. Enhancing Profits & Reducing Losses by Managing Material Environmental, Social & Governance Factors. In Academy of Management Proceedings (Vol. 2020, No. 1, p. 20390). Briarcliff Manor, NY 10510: Academy of Management.

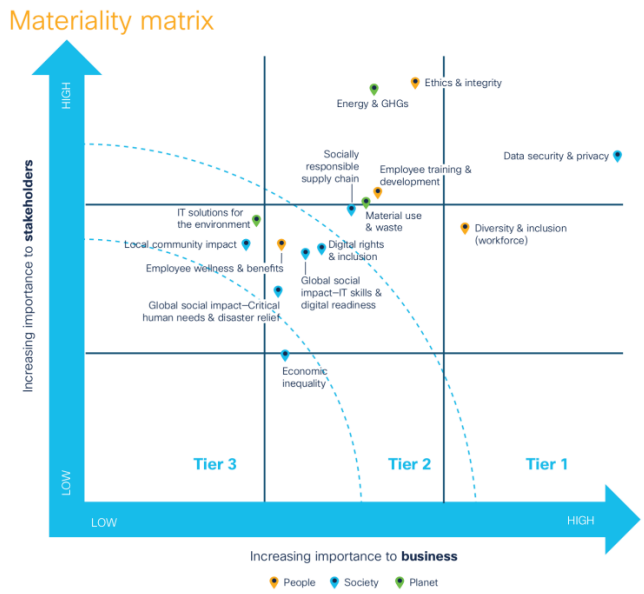
Serafeim, G. and Yoon, A., 2022. Which Corporate ESG News does the Market React to? Financial Analyst Journal.

Spandel, T., Schiemann, F. and Hoepner, A.G., 2020. Capital market effects of ESG materiality standards. Available at SSRN.

Welch, K. and Yoon, A., 2021. Do High-Ability Managers Choose ESG Projects that Create Shareholder Value? Evidence from Employee Opinions. Forthcoming, Review of Accounting Studies.

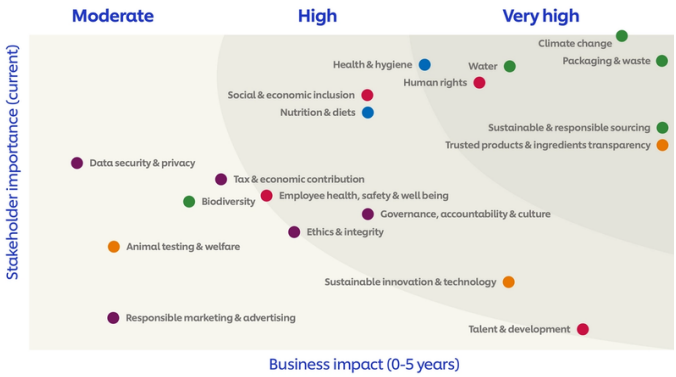
Figure 1 Example of a Materiality Matrix

Panel A



This figure presents the materiality matrix from Cisco, which is an excerpt from their 2019 Sustainability Report (https://blogs.cisco.com/csr/materiality-determining-what-matters-most).

Panel B



This figure presents the materiality matrix from Unilever, which is an excerpt from their Sustainability Reporting Center website (https://www.unilever.com/planet-and-society/sustainability-reporting-centre/our-material-issues/).