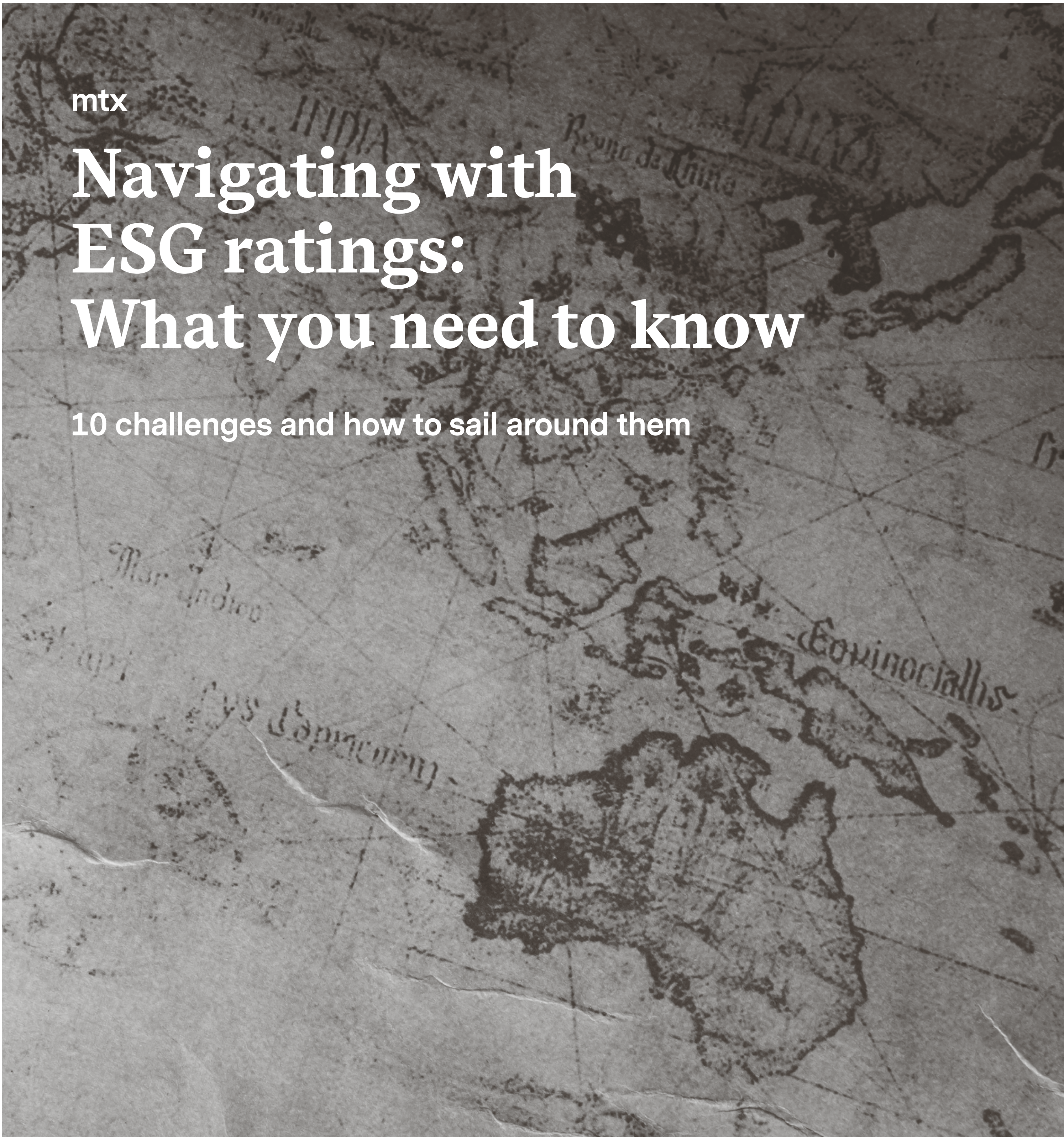
mtx
by Vontobel Asset Management

mtx

Navigating with ESG ratings: What you need to know

10 challenges and how to sail around them



Approved for institutional investors in: AT, CH, DE, ES, FI, FR, GB, IT, LI, LU, NL, NO, PT, SE, SG (Professional Investors only)

Lara Kesterton joined Vontobel Asset Management in June 2017 as an ESG Analyst for the mtx franchise within the Sustainable Equities boutique.

Prior to joining Vontobel, Lara Kesterton worked on developing decarbonizing pathways for international hotel groups, and on a World Bank project with the Maldives government. Lara has also worked with a number of private equity groups on renewable energy investments.

Lara is a lawyer by professional background, with five years' practice in a leading London law firm, specializing in banking and international finance.

Lara Kesterton attained her undergraduate degree at the University of Durham and a Master of Science in Environmental Change & Management from Oxford University.

Editorial

The first humans to cross an ocean were Polynesians who traveled the Indo-Pacific region starting at around 3000 BC. What makes this achievement even more noteworthy is the fact that those sailors managed to find their way without any maps. It took millennia for people from the Western world to match this great accomplishment, but this time they were equipped with cartographic material to navigate. However, the cartography back then was still a young discipline: the different schools were still evolving their approach to map projections, generalization (reducing irrelevant complexity) and design (a meaningful map must fit the audience's needs). Scholars now agree that cultural and social influences dominated early map making.

ESG investors might also think of themselves as early sailors lacking a definitive map, as investing according to ESG principles, which considers environmental, social and governance criteria, has not developed a common way to view the world (yet). Many ESG rating agencies are giving advice on how to navigate emerging territories, yet it is difficult to agree on a common mapping system.

Recent research into the challenges of ESG ratings have highlighted the disagreements among raters. In this white paper, we look into the reasons why ESG raters cannot agree and why some of these challenges are here to stay. Our goal with this publication is not only to caution against relying on a simple final score from an ESG agency for investment decisions, but also to offer a solution to the problems that investors face when they want to consider ESG criteria. We believe it requires a nuanced approach with a focused, multi-layered approach that helps you to see both – the important details and the bird's eye view of your investable universe.

Lara Kesterton
ESG Analyst

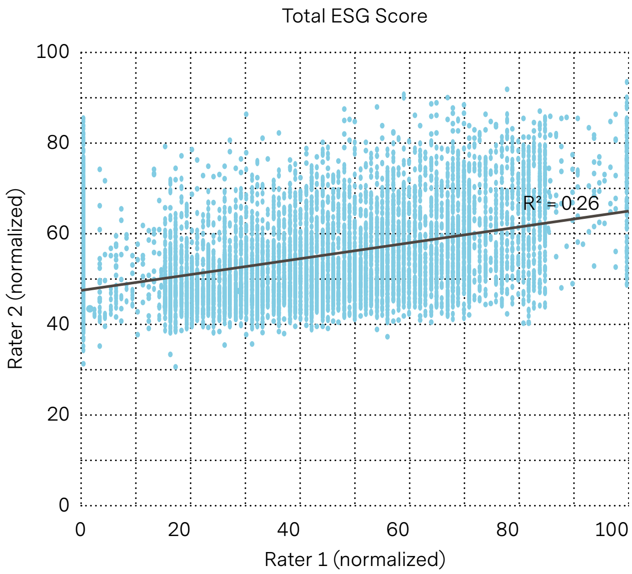
ESG ratings: Why can't raters agree?

Sustainable finance, after years of advocacy to become mainstream, is now growing significantly. According to one measurement, at the end of 2018 there were already some 18 trillion US dollars invested according to ESG integration approaches, an increase of 69% versus the end of 2016.[1]

With this tailwind, rating agencies that assess ESG factors to help investors make informed decisions on sustainable investing are booming, with more than 125 different agencies established world-wide.[2] These raters assess a number of different metrics, adding their own proprietorial magic for how to aggregate, weight, and come up with an overall number or grade. Akin to a credit rating score, this might give the impression of a consensus-drawn evaluation derived from hard facts and defensible figures, but these grades mask layers of subjectivity and hidden biases. In fact, approaches, and therefore results, of ESG raters differ widely as chart 1 illustrates.

Recent academic research performed similar analysis more broadly, finding a correlation coefficient of around 0.49[3] when comparing the scores of different leading ESG raters. To put this into context, this is contrasting to a coefficient of 0.96[4] (indicating strong agreement) for credit rating agencies, where of course the industry landscape and approaches are much more consolidated, also because of the longer history of such ratings. The research confirms that ESG rating agencies neither agree on what constitutes good ESG practice nor who is good or bad at it. Particularly, there was a stark disagreement in the tails of the ratings (very good and very bad companies), which is notable as many investors use these results to create best-in-class portfolios or avoid worst-in-class performers.

Chart 1: Can you spot the correlation?
Comparison of ESG scores between two leading ESG rating agencies



Source: Vontobel Asset Management as of November 15, 2019.
Company universe based on rater 1 universe.

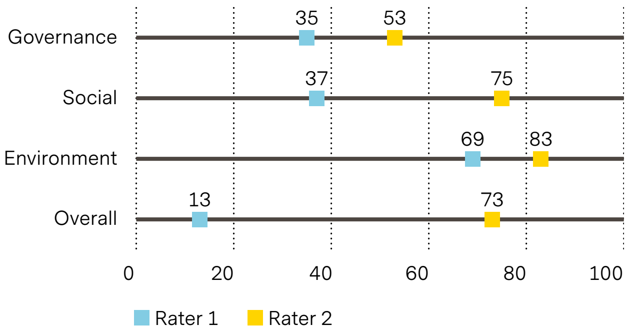
[1] Voorhes, 2018.
[2] Voorhes, 2018.
[3] This is the average of the mean correlation of the following four papers. Bender, et al., 2018 found correlation between four leading raters ranged from 0.47 to 0.76 with an average of 0.59. Gibson, et al., 2019 found average correlation between six prominent raters was 0.46. Berg, et al., 2019 found a correlation range of 0.42 to 0.73 with an average of 0.61 in their assessment of five leading ESG raters. Chatterji, et al., 2016 had the lowest mean correlation of 0.3 for six well-known raters (with a range from –.012 [indicating severe disagreement] to 0.67, and only a quarter of the correlations were higher than 0.5).
[4] Berg, et al., 2019.

> **"The research confirms that ESG rating agencies do not agree on what constitutes good ESG practice."**

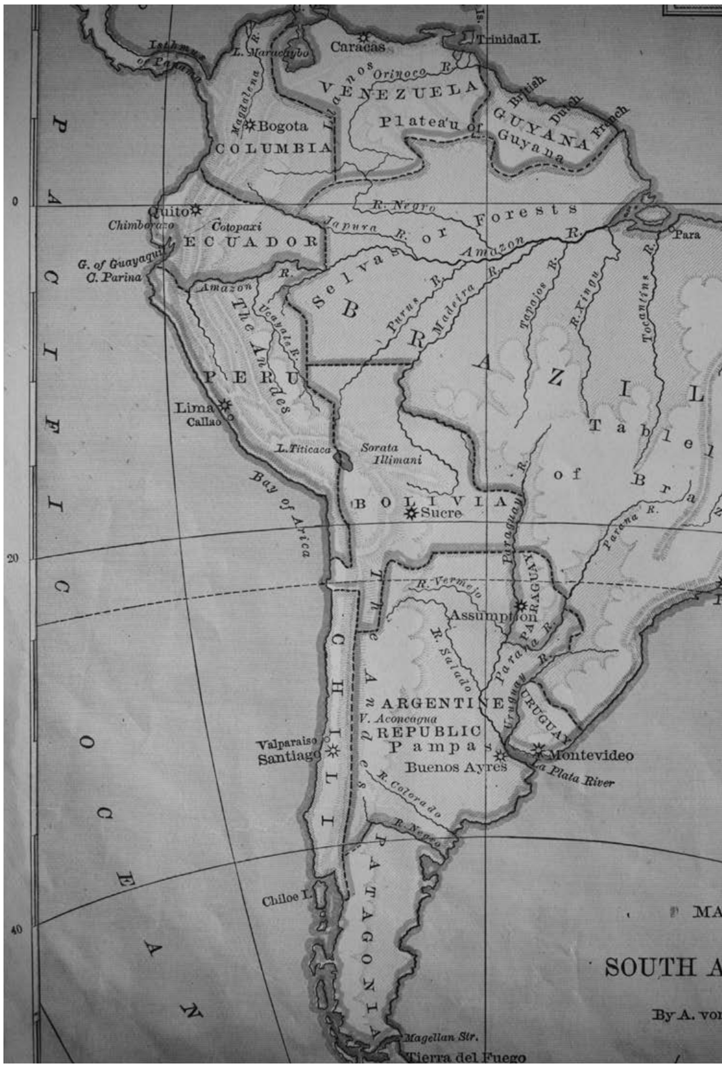
Why can't raters agree?
One underlying problem is that ESG raters serve various responsible investing interests (see our white paper ***Navigating ESG***[5] for the reasons for ESG investing and how to find the right ESG approach for your beliefs, and our white paper ***Evolution of Sustainable Investing and the case for integration***[6] for deeper background on ESG investment strategies). In practice, the raters usually go about the rating process by developing proprietary methodologies to rank and score companies on the panoply of ESG issues.

As input, ESG raters take data from multiple different sources and languages and use models to clean, organize, and weight these diverse data points to create comparability and to flag risks. As highlighted, for example, in chart 2, this can lead to different outcomes depending on who you ask.

Chart 2: How Toshiba Corporation is rated by leading rating agencies



Source: Vontobel Asset Management, as of November 15, 2019.



[5] Plinke & Münstermann, 2019.
[6] Hammerich & Kesterton, 2018.

Ten challenges of ESG ratings

The scoring models used by ESG raters of course have their merits by giving structure to decision making, but they also are at risk of giving the impression of scientific rigor, when in fact ESG practice is still an art. In the case of ESG ratings, they come with many challenges.

Table 1: Top 10 challenges of ESG ratings

CHALLENGE	EXPLANATION
Material factors	What ESG topics are looked into? What is considered a material issue?
Measurement	What metrics are scored for these material issues?
Data quality	What data sources are used for the metrics? How reliable are they?
Gaps treatment	How are data gaps treated? Penalized? Filled with averages?
Timing aspects	How often do raters rate? Reporting lag and backward looking data concerns
Rater bias	Raters' world view has latent influence on how metrics are interpreted
Weighting methodology	How are metrics aggregated into a score?
Controversy handling	What relevance/red-flag importance is given to controversies?
Benchmarking	Is the final rating based on a relative or absolute scoring?
Aggregation of ratings	Fund average score gives a false impression of wide score divergence

Source: Vontobel AM

Looking into the ten most prominent challenges, listed in table 1 in more detail, reveals the complexities when trying to capture the real world into a scoring model.

1. **Material factors**
   Considers what ESG topics should be included in the model, e.g., while greenhouse gas emissions will be commonly assessed, indigenous rights, employee organizations, or lobbying might be more niche topics for assessment and only scored by a few. The number of data points evaluated by raters vary from 10–>400, although there is good evidence that counting too much merely weakens the real signal aimed for.[7]

2. **Measurement**
   Raters use different metrics to evaluate a topic, e.g., to evaluate employee health and safety, raters choose from 20 different data points to score this topic.[8] Some research found this to be the dominant reason for rater divergence.[9] Peeling back the layers of what gets measured, the raw underlying data is more inconsistent than you might think.

3. **Data quality**
   Related questions are: how defensible is the ESG data? Is it pure marketing information, as non-financial information is not required to be certifiable or defensible in the same way that financial statements are? Frequently, metrics supplied by companies are patchy, inherently backward looking, and tend to fall into "good news" storytelling. Some raters exclude data provided by the company itself, while this can naturally be a rich data source. Similarly, as ESG metrics are frequently qualitative, raters must choose how they interpret and score descriptive matters.

[7] The Sustainability Accounting Standards Board (SASB) is leading the charge on addressing this with its endeavor to create consensus on material ESG issues for each industry and sub-sector.
[8] Kotsantonis & Serafeim, 2019.
[9] Berg, et al., 2019, Chatterji, et al., 2016.

4. **Gaps treatment**
   It is common for companies not to report on all indicators (let alone provide industry comparable metrics). Different statistical tools can be used to fill the gaps with widely different outcomes.[10] Interestingly, a few studies found larger firms experience more disagreement in their scores suggesting again that more data points can lead to more disagreement between raters. An active investor with good relations with the firm can sometimes overcome data gaps by direct dialogue.

5. **Timing aspects**
   The frequency with which raters evaluate a company can have a material bearing on discrepancies between scores. An annual review is not uncommon, but also time gaps of two years between the latest updates of different raters may exist.

6. **Rater bias**
   The rating houses have a natural (sometimes outspoken) slant, e.g., a focus on best-in-class, risk, momentum, and climate. It has been observed that raters based in civil-law countries (e.g., Germany and France) are more focused on social issues, whereas common-law countries (e.g., the UK and US) have a shareholder-centric approach and therefore have higher focus on governance issues.[11] In addition to explicit biases (which are reflected in the materiality assessment), research has shown an unexplained or unconscious "rater effect", in that when a rater is generally positive (or negative) on a company this is reflected across the board, including on unconnected indicators. This could account for 14–18% of rater disagreement.[12]

7. **Weighting methodology**
   Next, raters need to assign how much importance to give an indicator in their model. This is largely subjective and not always transparent. Most models have indicators with little to no statistical significance – meaning they are being scored without having any real impact on the overall ESG score (or any link to financial performance).[13]

8. **Controversy handling**
   Controversy handling is the walk of the sustainability talk, and for many raters they have a high prominence in scoring. To be comparable, controversial incidents have to be evaluated for impact on society and for the business – once again an open field for subjectivity and disagreement.

"An aggregated score is even further removed from the underlying raw data."

9. **Benchmarking**
   As the rater translates the scoring into a final rating, an important input is also the perspective taken.

   Relative scoring is commonly used to benchmark performance against peers. But this raises the question – what is the right peer group? Universal comparisons or against the industry peers (there are merits for both)? If the latter, again, raters choose from different industry classification systems, such as GICS, BICS, IVA industries, or perhaps an in-house division of industries. Then throw in to the mix how to treat diversified companies, and no wonder a leader in one classification can be only average in another rater's eyes. Additionally, relative scoring can of course miss the point on sustainability if the entire industry is not addressing the issue well enough.

   Absolute scoring is the alternative approach and scores on preset ranges or optimal levels. Subjectivity creeps in on who sets the benchmark and then this leads to natural tilts away from certain industries or countries, which commonly underperform in certain areas, e.g., diversity in the financial sector or on Chinese boards.

10. **Aggregation of ratings**
    Portfolios are also scored on their average ESG rating. In truth, the average fund scores tend to be tightly clustered in a narrow spread, therefore, a top-rated fund may not have an average score notably ahead of a weak fund. At this fund level the aggregated score is even further removed from the underlying raw data and are now in black-box territory in terms of what the scores really ought to tell you – how exposed you are to risks and whether those risks have been adequately priced in.

[10] E.g. do you assign the industry average (or universal or home market peer group average) or score with lowest score or use some other statistical model or not score at all? Kotsantonis & Serafeim, 2019 examines this in detail.
[11] Gibson, et al., 2019.
[12] Berg, et al., 2019.
[13] Berg, et al., 2019.

How to sail around these challenges?

A deafening demand across the ESG industry is for companies to supply better quality and more comparable data. This should address a major reason for disagreement amongst raters. There are various voluntary industry and legal initiatives[14] working to create a common set of metrics on which all companies should report on.

Another way to mitigate the problem, a new wave of artificial-intelligence-driven ESG ratings are being designed to overcome human unconscious biases and normalize for size and industry skews. Other major trends are increasing use of unconventional data sources[15] to get more impartial risk insights as well as consolidation within the rating industry. The major raters have been on a land grab in the last few years buying up smaller, niche players, suggesting a consolidation on ESG theorization may emerge. However, at the same time, sell-side analysts have entered the space adding alternative views.[16]

An active, high-conviction manager should look beyond aggregated ratings

For the thoughtful investor, this disillusion with ratings requires looking beyond frameworks and adopting a multi-layered approach. To start with, use informative data from the ESG raters to feed an own in-depth assessment to enrich fundamental equity analysis. A step-by-step process of investigation leads to a much more detailed and holistic understanding of a company: its flaws and beauty spots but always focusing the few issues that are really material to that company. This detailed appreciation of the top ESG risks that can impact performance is much more informative to an active investor than the specific score crunched out at the end of the rater's model. The real goal is to use ESG information to understand if the company in question has the ability to withstand its top risks in a one-to-five-year time frame.

> **"The hunt for high ESG ratings does not result in outperformance."**

Still, at some point you want to aggregate your findings on a portfolio level and this is when you have to make sure to not lose details when zooming out. One way to go about it is to visualize the findings on a stock level in a tile chart as shown in chart 3.

The chart illustrates an assessment of the exposure of a portfolio of stocks to key environmental risks, broken down by industry sector. This is an aggregation of the more detailed company-by-company ESG risk assessment. This way, risk concentrations are easy to spot, without losing the important details on where exactly those risks come from.

Conclusion: There is no standard ESG methodology in the market – will there ever be?

As recent research notes, the inconsistency in ratings "(does) not discredit ESG data or the practice of scoring… it underscores the danger of relying on a simple final score for investment decisions."[17] In particular, the hunt for high ESG ratings does not result in outperformance, and does not necessarily even mean you are maximizing the sustainability of your investments. At the end, the ESG investment methodology should reflect the responsible investment approach that the investor is seeking.

[14] EU Non-Financial Reporting Directive has required ~6,000 EU companies to publish ESG data since 2017 annual results. Plenty of other regulatory requirements come from stock exchanges (UNSSE, ESMA); international and domestic law (e.g. legislation in discussion under EU Action Plan, French Article 173, China mandatory ESG disclosure by 2020); principles frameworks (i.e. ICMM, TCFD, SDGs, GRI, UN Global Compact); or voluntary disclosure frameworks (SASB, GRI, CDSB). The alphabet soup is discussed further in Temple-West, 2019.

[15] E.g., geographic information systems data (e.g., for real estate at risk), loyalty scores and customer reviews, independent product recall data, supply chain mapping, non-government organization reports, employee review sites and many more.

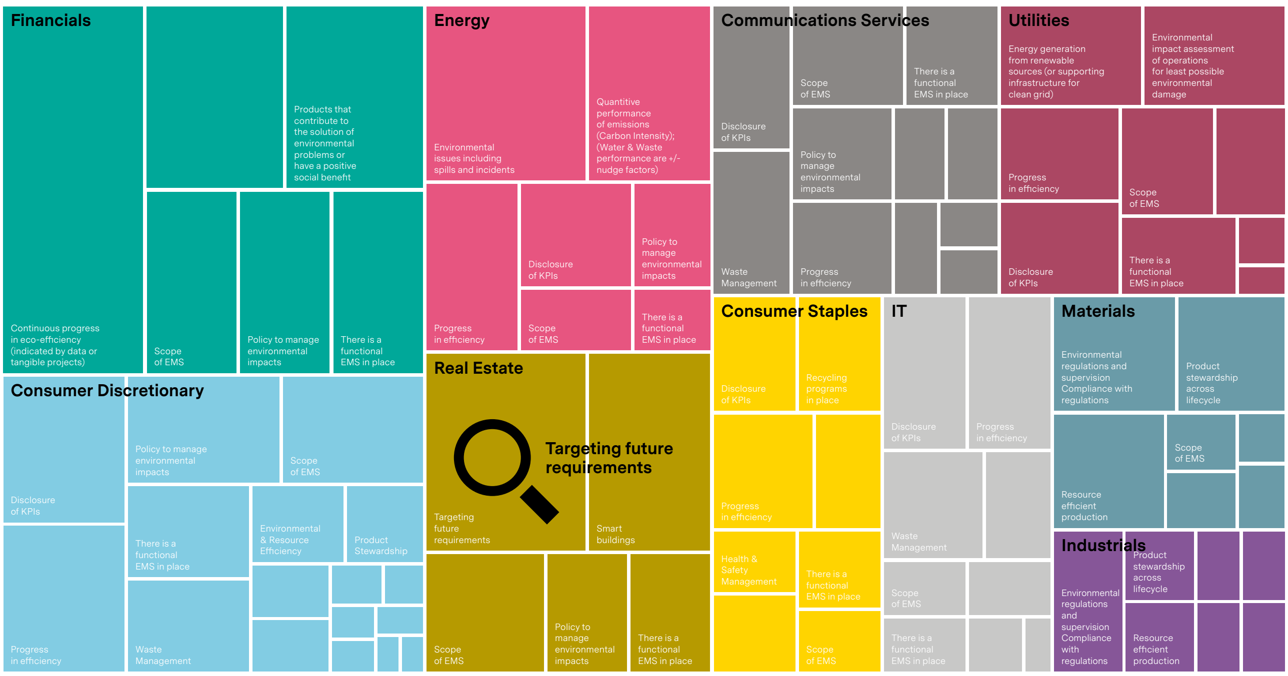
[16] Naumann, 2019.

[17] Yonts, et al., 2018, p.9.

For us as an active, high-conviction equity manager, mainly active in emerging markets, this means we conduct our own deep dive ESG analysis, in particular for companies that are not fully covered by ESG raters. We prefer an absolute perspective, setting a minimum standard to make a company investable. We put a lot of focus on controversies, which might result in a company becoming non-investable even if it passes on the average of scores. Ultimately, we concentrate on the most important risk areas to achieve a more holistic conviction on how exposed a company is to ESG factors and how well prepared it is to navigate these challenges.

The complexity of the real world issues being evaluated from environmental, social and governance perspectives, and the difference in objectives of ESG investors, may mean raters can never achieve a robust, consensus view in the same way as credit rating houses, for example. A better analogy is the diversity of opinions of financial analysts on the sell side, even though derived from standardized financial data. While this may make decisions for investors more difficult, it also offers opportunities for those able and willing to appreciate the intricacies involved with ESG assessments.

Chart 3: Example: Aggregation of environmental risks on a portfolio level, by industry sector



Source: Vontobel. Data is for Vontobel Fund – mtx Sustainable Emerging Markets Leaders, as of June 28 2019. The bigger the size of the tile, the larger the aggregated risk to that issue. Single-stock risks are weighted by portfolio holdings.

Navigating with ESG ratings

The wide ocean

18
Trillion USD
Assets under management in ESG integration approaches

69 %
Asset growth in the previous 2 years

Challenges during the voyage

>125
Raters
Wide set of service providers

10 – 400
Data points
Evaluated by an ESG rater

0.49
Correlation coefficent
Mean correlation of scores between ESG raters

Compare this with smooth sailing for credit analysts:

0.96
Correlation coefficent
Correlation of scores between credit raters

Sources: Navigating with ESG ratings: What you need to know, Vontobel, as of Nov 2019.

Sheltered bays

New frameworks

An alphabet soup of initiatives is working on action plans for better ESG clarity.

Artificial intelligence

Big data analysis, real-time insights, can remove human biases.

Alternative data sets

From satellite flood data to blockchain mapping of supply chains, alternative, impartial data sources are becoming critical risk signals.

Consolidation of raters

Significant mergers & acquisition in the ESG rating agency industry increases key player concentration and more homogeneity of ratings.

Solution

Adopting a multi-layered approach, no reliance on simple final scores

List of Sources

Bender, J. et al., 2018. A Blueprint for Integrating ESG into Equity Portfolios. The Journal of Investment Management, 16(1).

Berg, F., Kölbel, J. F. & Rigobon, R., 2019. Aggregate Confusion: The Divergence of ESG Ratings. MIT Sloan Research Paper, 5822(19).

Chatterji, A. K., Durand, R. & Touboul, S., 2016. Do ratings of firms converge? Implications for managers, investors and strategy researchers. Strategic Management Journal, 37(8), pp. 1597-1614.

Gibson, R., Krueger, P., Riand, N. & Steffen Schmidt, P., 2019. ESG Rating Disagreement and Stock Returns. [Online]
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3433728
[Accessed 18 11 2019].

Hammerich, C. & Kesterton, L., 2018. Evolution of Sustainable Investing and the case for integration. [Online]
Available at: https://am.vontobel.com/en/insights/evolution-of-sustainable-investing-and-the-case-for-integration
[Accessed 18 11 2019].

Kotsantonis, S. & Serafeim, G., 2019. Four Things No One Will Tell You About ESG Data. Journal of Applied Corporate Finance, 31(2), pp. 50-58.

Naumann, B., 2019. Credit rating agencies join battle for ESG supremacy. [Online]
Available at: https://www.ft.com/content/59f60306-d671-11e9-8367-807ebd53ab77
[Accessed 18 11 2019].

Plinke, E. & Münstermann, L., 2019. Navigating ESG. [Online]
Available at: https://am.vontobel.com/en/insights/navigating-esg
[Accessed 18 11 2019].

Temple-West, P., 2019. Companies struggle to digest 'alphabet soup' of ESG arbiters. [Online]
Available at: https://www.ft.com/content/b9bdd50c-f669-3f9c-a5f4-c2cf531a35b5
[Accessed 18 11 2019].

Yonts, C., Allen, J. & Zhou, M., 2018. CG Watch 2018. [Online]
Available at: https://www.clsa.com/wp-content/uploads/2018/12/CG-Watch-2018-Short-version-181205-L.pdf
[Accessed 18 11 2019].

Your point of contact

Marc Bindschädler
Client Portfolio Manager
Vontobel Asset Management

marc.bindschaedler@vontobel.com
T +41 58 283 58 87

Lara Kesterton
ESG Analyst
Vontobel Asset Management

lara.kesterton@vontobel.com
T +41 58 283 68 54

Imprint

Editor
Vontobel Asset Management AG
Gotthardstrasse 43
8022 Zurich, Switzerland

Text
Jens Finke, CFA

Design
MetaDesign AG

Creation & Realization
Design Team, Vontobel

Images
Gettyimages
Vontobel

Vontobel Asset Management AG
Gotthardstrasse 43
8022 Zurich
Switzerland

vontobel.com/am