ESG outperformance: Not about one factor

Far from country bets, crowding or overfitting, so-called “dark green” ESG indices have displayed out-performance across regions and over time. These indices have also exhibited clear signs of risk reduction in volatile market conditions. While the focus on higher ESG ratings at an index levels has delivered excess returns especially in Emerging Markets, which fits nicely with academic theory, many of the drivers of outperformance have tended to be more structural.

Looking at “dark green” ESG index rules, there are two main construction pillars: exclusions, based on activity involvement and controversies and inclusions, based on ESG rating classification criteria. We consider these two pillars as complementary generators of potential excess returns. This view is also widely supported by academic research. The exclusion approach provides more macroeconomic and single-stock stability thanks to structural shifts in the benchmark while the inclusion pillar generates financial outperformance, notably on the back of shifts in the style footprint. Given these different ESG dimensions, it is therefore difficult to pinpoint ESG outperformance to simply “one factor”.

With the benefit of hindsight, the resilience of ESG indices during the COVID-19 crisis has only been a continuation, and to some extent an acceleration, of an existing dynamic at play. However, many investors and practitioners are divided as to whether the performance of ESG indices is due to herding effects, to sector biases, avoidance of tail risks or actually to the financial strengths of higher ESG rated companies. This paper uses performance and risk attribution techniques over the 2015 to 2020 period. The analysis is based on historical index holdings and ESG data which are used to examine the causality of potential sources of outperformance of ESG indices in the Equities space. We selected the MSCI Low Carbon SRI Leaders indices (thereafter, “LCSL”) as a representative “dark-green” ESG index, which in our view, strikes a balance between broadly-inclusive, low tracking error indices with minimal exclusion rates in controversial activities and fairly exclusive, bottom-up indices with high concentration/ tracking error risk and reduced concept portability in single countries.

ESG performance: a behavioural analysis

The COVID-19 crisis stands out since it captures environmental, social, and governance aspects alike. While LCSL ESG indices were not spared from sizeable drawdowns in 2020, they proved to be considerably more resilient than their non-ESG counterparts. In fact, we find strong evidence of a superior risk return trade-off for ESG indices – an important driver for ESG adoption among less ESG-savvy investors. For all markets covered in this paper we find evidence of solid downside protection: COVID-19 related drawdowns were 30bps and 90bps less for the LCSL versions of MSCI World and EM, respectively. For Japan (290bps) and Europe (210bps) we find even stronger evidence.

Nevertheless, when looking at year to date performance for ESG indices vs. their non-ESG counterparts, it could be seen that drawdown reduction alone does not explain the
full performance picture. In fact as of the end of November 2020, year-to-date performance differentials between ESG and non-ESG indices are larger than the drawdown differences might suggest – pointing to a more pronounced recovery from ESG indices. Looking at the last five years, a timeframe across which ESG quality is comprehensive and reliable, it could be seen that both elements, namely risk reduction in periods of drawdown, and wider equity market outperformance are not unusual. Rather 2020 marked a continuation of existing trends which ESG strategies had displayed during the bull-market leading up the COVID-crash1.

A key driver behind the strong performance has been the ability of ESG strategies to capture market upside, particularly over the last five years – a result that at first might be difficult to reconcile with a quality tilt ESG strategies are often said to feature. Figure 1 summarises the relative performance of the ESG strategy across different market environments. It is worth pointing out that all five ESG strategies under review feature positive upside capture over the last five years.

We also find evidence of a growing role of financial outperformance across the index members (i.e. cross-sectional, rather than across time). For World LCSL we find that 51% of index constituents display either the above discussed financial outperformance or risk reduction characteristic. For EM LCSL this share increases to 57%. Specifically, for 29% of index constituents, the ESG methodology resulted in an overweight of the constituent and this overweight translated into financial outperformance. Similarly for 28% of index constituents’ underweights paid off by reducing the exposure to underperformers.

These phenomena have given rise to several theories circulating around the investor community to explaining recent ESG performance. Before we explain our multi-factor attribution results, we aim to debunk three particularly entrenched beliefs, namely (i) the role of inflows, (ii) the role of country allocation and (iii) the role of factor tilts.

While the ESG investing marketplace is nascent, the investment style is growing rapidly. In fact, in the last 12 months alone, UCITS equity ESG ETFs have attracted over EUR 60bn in inflows, a threefold increase in inflows over the prior year. And since, as always, ETF flows are only an indicator of broad shifts in the investment community, investor allocation towards ESG is also rising rapidly. Some might even be under the impression that, given exclusion rates of 50% and above, capacity in ESG indices may naturally suffer from bottlenecks. It could be therefore worth asking the question, are recent inflows into a narrow ESG universe the key driver of ESG outperformance? We find no significant evidence of such a conclusion. In fact, when regressing weekly excess returns on weekly flows, any association between the two appears spurious (correlation = 13%). Also we find no association between flows and current valuation premiums over MSCI World. On a P/E basis the World LCSL index continues to trade at a discount vs. MSCI World. In fact over the 12 months, the period of strongest outperformance and inflows this valuation gap has widened further2.

It is also worth keeping in mind that ESG outperformance is not the result of data mining or over-fitting in the index construction nor ESG data cherry picking. In fact, when cross validating the replication of different ESG objectives using

---

1 For example, when comparing the Sharpe ratio across World, its major building block, Europe, US and Japan, as well as Emerging Markets, we find that ESG strategies provide superior risk-return combination across time and across regions compared to their non ESG index across the 5Y horizon.

2 While the P/E discount of ESG strategies might come as a surprise to many investors, ESG strategies do trade at a price-to-book and price-to-sales premium against MSCI World, in line with a quality bias in these strategies.
DWS proprietary engine data points and applying simple sector/region neutral index construction approaches, we can recreate similar levels of ESG outperformance as those seen in traditional MSCI ESG indices.

### Attribution: a deeper dive challenges conventional wisdom

Multi-factor attribution results contradict established beliefs and provides further perspective than one-dimensional attribution results

The last but not least common misconception is the reduction of ESG performance to a single set of drivers, such as currency, country biases, growth, large cap biases. Our multi-factor attribution results indeed confirm that such effects are in fact broadly negligible both from a risk and from an excess return perspective when compared with other factors.

The following provides an overview of the multi-factor attribution results. In this part, we use traditional one-dimensional and more elaborate multi-factor attribution techniques to identify the drivers of outperformance in the main regions of focus for European investors: World (which is composed of two thirds of US stocks currently), Europe and Emerging Markets.

The figure below offers a 5-year attribution picture for these indices, broken down in performance drivers: country, currency, sector, factor (also called style) and non-factor residual, called selection from an attribution perspective. The second dimension on this figure is the average contribution, from a risk perspective, of each such driver to the tracking error, which is the standard deviation of return differences between ESG and respective non-ESG counterpart. This dimension enables to estimate, for example, whether the smaller 5-Y attribution component of a driver is fortuitous (as such a component has a high inherent active risk component) or is due to the fact that such component has indeed little impact on overall excess returns. In the figure below, it is clear that:

- Currency risks, and to a certain extent country risks in developed markets, have broadly little relevance in the overall excess returns
- The currently high sector driven component of the performance attribution, mainly originating from the US, is not necessarily likely to reproduce (more than one sigma when comparing annualized excess return contribution with risk contribution taken as a proxy of standard deviation of excess returns (sigma))
- The relatively low level of the selection effect may be temporary given the high contribution to tracking error they typically have.
- As a direct consequence, trying to reduce other biases than idiosyncratic (selection) biases would further increase the potential effects of selection, to a potentially excessive level
- Country effects in EM are likely to be more important than sector biases
- The high selection effect over 5Y in emerging markets is not a statistical outlier (less than one sigma)

And most importantly: ESG performance drivers are numerous. The long-term outperformance tends to be a culmination of sector, factor and country factors. Reliance on these key pillars of ESG outperformance provides investors with outperformance with just a moderate tracking error. A key aim of our analysis in the following section will be to assess the extent to which they are integral parts of a “dark green” ESG approach.

![FIGURE 3. CUMULATIVE ATTRIBUTION AND FACTOR CONTRIBUTION TO EXCESS RETURNS IN % (Y AXIS) VERSUS AVERAGE ACTIVE RISK CONTRIBUTION OVER THE PAST 3 YEARS](image)

Nevertheless, the respective contribution from these drivers has not been constant across time. Figure 4 confirms that the historical contribution of each driver (and its subcomponents) varied across time and where sector biases, style biases and selection effects do not move in sync. We find there is a close to zero correlation between these performance drivers, and only 25% correlation between energy underweight effects and technology overweight effects. This model and decomposes the active return, i.e. the return of a portfolio (ESG index) over the return of its benchmark (non-ESG parent), into an allocation effect and selection effect. The factor-based approach, on the other hand, decomposes the active return into the contribution of a series of factors (generally speaking, for equities these include the market, country, industry and style factors).

---

2 Also MSCI has conducted extensive research into replicability of outperformance and risk reduction using different shades of ESG strategies (MSCI, 2019).

4 Specifically, all performance attribution is based on Bloomberg PORT modelling. The two underlying models used are broadly based on (i) Brinson Hood and Beebower (1986), and (ii) the factor-based approach proposed by Grinold and Kahn (1999). The former is often referred to as the Brinson model and decomposes the active return, i.e. the return of a portfolio (ESG index) over the return of its benchmark (non-ESG parent), into an allocation effect and selection effect. The factor-based approach, on the other hand, decomposes the active return into the contribution of a series of factors (generally speaking, for equities these include the market, country, industry and style factors).

---

Sources: DWS International GmbH, Bloomberg, as of November 2020. Past performance is not a reliable indicator of future returns. Forecasts are based on assumptions, estimates, views and hypothetical models or analyses, which might prove inaccurate or incorrect.
points to very diverse potential sources of excess returns; it is also worth noting that the negative selection effects are rather recent and only started in the wake of the COVID-19 crisis. These results are also representative of the general situation in US equities. Indeed Figures 5 and 6 show indeed more persistent performance drivers in Europe and EM, with style being a persistent performance driver in Europe and selection (albeit more volatile) in Emerging Markets.

But what we believe is most striking and contradicts conventional wisdom among investors regarding ESG investing is that the outperformance of ESG does not always originate from selection effects, but rather it is generated currently by “ordinary” sector biases, and also style biases in developed markets. We will explain that those biases are necessary and form an integral part of an ESG approach. We will also show that those results do not contradict previous academic and practitioner analysis, and that such results are partial: selection effects after applying a one-dimensional, Brinson attribution analysis by sectors will entail the effect of style biases and will appear positive over most time periods.

FIGURE 4. WORLD – HISTORICAL CONTRIBUTION TO EXCESS RETURNS OF THE MAIN PERFORMANCE DRIVERS
Based on MSCI World Low Carbon SRI Leaders index vs. MSCI World index, 30/09/15 – 30/09/20

Source: Bloomberg, DWS International GmbH. As of 30/09/2020. All returns calculated in USD. Relative performance rebased to 0 at 30/09/2015. Past performance, actual or simulated, is not a reliable indicator of future results.

FIGURE 5. EUROPE – HISTORICAL CONTRIBUTION TO EXCESS RETURNS OF THE MAIN PERFORMANCE DRIVERS
Based on MSCI Europe Low Carbon SRI Leaders index vs. MSCI Europe index, 30/09/15 – 30/09/20

Source: Bloomberg, DWS International GmbH. As of 30/09/2020. All returns calculated in USD. Relative performance rebased to 0 at 30/09/2015. Past performance, actual or simulated, is not a reliable indicator of future results.
In the remainder of this section, we focus on the most prominent performance drivers identified: Sector, style and selection effects.

Sector attribution: Technology and energy most prominent drivers in the US, less so in Europe or EM

Here, we return to the earlier discussed “one-dimensional” analysis. Figure 7 below is a traditional attribution analysis per GICS sector. It confirms the roles of the IT overweights and Energy on the allocation side and points to positive selection within financials and industrials, which may be linked to ESG selection. Outside of this analysis, results are slightly inconsistent depending which sector classification is used (selection effect goes from -1.4% to +2.0% when using the ICB classification instead of GICS). The results also show only mildly negative selection effects, which would have been positive had the analysis be undertaken in May 2020 or over a shorter period of time, as many papers suggest. Here again, those selection effects only point to the residuals after adjusting for sectors and masks the contribution of, amongst others, style effects. Similar analysis in Europe and EM shows positive selection effects. In none of these two regions is any material sector allocation effect worth mentioning: the sector allocation effect is mainly concentrated in the US, where the IT and Energy sectors dominate and performance contribution has been steady over the 5-year period studied.

**FIGURE 7. WORLD ESG: 5Y CUMULATIVE SECTOR ATTRIBUTION VS BENCHMARK (09-2015 TO 09-2020)**

<table>
<thead>
<tr>
<th>Ranking based on total attribution</th>
<th>Over/underweight (%)</th>
<th>Total Attribution (Alpha) (%)</th>
<th>Selection Effect (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information Technology</td>
<td>-2.73</td>
<td>3.01</td>
<td>3.30</td>
</tr>
<tr>
<td>Industrials</td>
<td>-2.46</td>
<td>-2.46</td>
<td>2.14</td>
</tr>
<tr>
<td>Energy</td>
<td>-2.07</td>
<td>2.64</td>
<td>1.31</td>
</tr>
<tr>
<td>Consumer Staples</td>
<td>-0.37</td>
<td>-0.46</td>
<td>-0.27</td>
</tr>
<tr>
<td>Utilities</td>
<td>-0.64</td>
<td>2.64</td>
<td>0.27</td>
</tr>
<tr>
<td>Financials</td>
<td>-0.48</td>
<td>-0.46</td>
<td>-0.46</td>
</tr>
<tr>
<td>Real Estate</td>
<td>-1.21</td>
<td>-1.64</td>
<td>-1.68</td>
</tr>
<tr>
<td>Communication Services</td>
<td>-2.5</td>
<td>-2.04</td>
<td>-1.79</td>
</tr>
<tr>
<td>Health Care</td>
<td>-0.63</td>
<td>-2.43</td>
<td>-2.26</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td>3.20%</td>
<td>-1.40%</td>
</tr>
</tbody>
</table>

Source: DWS International GmbH, as of 30/09/2020, based on historic index holdings for MSCI Low Carbon SRI Leaders Series, all performance attribution based on Bloomberg PORT. Past performance, actual or simulated, is not a reliable indicator of future performance. Total attribution = selection + allocation + currency effects.
Factor attribution: the quality bias is consistent and produced outperformance

In this section, we dive deeper into the contribution of factors to performance using the MAC2 global equities risk model of Bloomberg PORT. These findings of our analysis point to two main phenomena:

1. A quality bias is present across all regional indices, more precisely lower leverage, higher profitability and higher earnings stability. In all regions considered, those biases did bring excess returns. This points to an inherent bias of an ESG approach which is recognised broadly in academia.

2. Certain factor biases have little aggregate exposure but relatively high contribution to performance: this is because they are present in different signs in certain segments of the index: it is the case for example regarding:
   a. Lower size in the US IT segment (due to controversies as detailed below)
   b. Low volatility in Healthcare and Industrials across DM regions
   c. High growth, momentum and higher volatility in Consumer Discretionary within Emerging Markets

In this situation and in the absence of a clear link towards ESG, those effects (around 0.6% for MSCI World LCSL over 5Y) could be added back to the selection component.

Selection effects: The known unknown in performance attribution?

While a significant part of ESG outperformance could be attributed to the above discussed sector and factor tilts, ESG performance attribution in our opinion would be incomplete without a discussion on selection effects. In our multi-factor analysis, the selection effect could be understood as the residual performance, unexplainable by the model. Given the model’s comprehensive coverage, selection could be best thought of as an idiosyncratic component. There is presently a general understanding that the outperformance effects due to an ESG approach could be read in the “selection” effects. Our analysis will rather show the contrary: not only are factor biases necessary by-products of an ESG approach but also more structural biases such as sector allocation or even certain large caps are necessary biases originating from a “dark green” ESG approach.

Looking at MSCI World LCSL, significant negative selection returns could only be found across three sectors, healthcare, consumer discretionary, and IT – all three among the top-performing sectors pre- and post-COVID-19. A detailed assessment of constituents shows that the majority of the negative selection effect is associated to the exclusion of selected mega caps in these sectors. In fact, six of the top 10 largest weights in MSCI World have not been part of the ESG index over the last five years at all. With an average annual return of over 20% for these names, mostly due to idiosyncratic success stories, little of this performance will be explainable via a multi-factor approach. In the case of MSCI World LCSL, factor (country + style + sector) contribution among the top 10 has in fact been flat, while the selection effect in the top 10 is about 350bps (410bps in top 25) and hence a highly significant share of overall selection effects. This skew in selection is particularly prominent in indices with larger performance dispersion, such as the US and EM, and lower in Europe and Japan.

Sources: DWS International GmbH, Bloomberg, as of November 2020. Past performance is not a reliable indicator of future returns. Forecasts are not a reliable indicator of future returns. Forecasts are based on assumptions, estimates, views and hypothetical models or analyses, which might prove inaccurate or incorrect.

5 We are aware of certain weaknesses of the MAC2 risk model of Bloomberg such as the non-consistency between factor contribution in the multi-factor space and single-dimensional attribution, which should be, to some extent, addressed in the next MAC iteration. We find however that the results delivered are consistent with research published by MSCI using MSCI Barra Bara risk model, November 2020.
At the same time, it is worth pointing out that the mega-cap effect also masks significantly more positive selection results for the vast majority of index weight. As the below chart illustrates, especially among the smaller names of the index with weights of less than 0.03% (accountable for 70% of overall index population), selection effects are positive. In Emerging Markets LCSL, see Figure 10, where selection effects are very consistent across stock sizes, and Europe LCSL indices we find further confirmation of this pattern.

Overall, ESG outperformance occurred not because of, but rather despite, mega-cap exposure. Should ESG investors be put off by the presence of negative selection effects? We argue that they should not and that selection is only a residual fraction of the effects of ESG. While past selection effects certainly illustrate some of the opportunity costs associated with a “dark green” ESG approach, they also illustrate that the directional impact of selection effects is likely to fluctuate over time. Furthermore, as section 4 of this paper highlights, selection effects are a voluntary trade-off in a “dark green” approach.

FIGURE 9. MSCI WORLD LCSL: SELECTION EFFECTS
Selection effect (orange/green, LHS) by size category, Category population (grey, RHS) in % of total number of stocks


FIGURE 10. MSCI EMERGING MARKETS LCSL: SELECTION EFFECTS
Selection effect (orange/green, LHS) by size category, Category population (grey, RHS) in % of total number of stocks

Linking ESG excess returns to index construction: the concept of necessary biases

The material role of comprehensive ESG index rules in practice

The chart below looks at a sample practical, “dark green” approach to ESG index construction:

<table>
<thead>
<tr>
<th>Benchmark kill rate</th>
<th>10%</th>
<th>24%</th>
<th>22%</th>
<th>-2.5 – 3x overweight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rationale</td>
<td>Avoid involvement in non socially responsible activities (e.g. carbon-intensive, tobacco/alcohol/weapon-related activities)</td>
<td>Avoid association with companies subject to lawsuits in violation of sustainability or social norms</td>
<td>Avoid companies with high environmental, social or governance related risks</td>
<td>Select ESG leaders in their sector/region category</td>
</tr>
</tbody>
</table>

*FIGURE 11. MAIN INDEX EXCLUSION AND INCLUSION RULES IN RELATION WITH ESG ASPECTS, RATIONALE AND RESULTING CONTRIBUTION TO ACTIVE SHARE IN THE ESG INDEX*

Source: DWS International GmbH, as of November 2020

It is clear from the figure above that a more conservative ESG approach is not limited to focusing on stronger ESG ratings. And these index construction pillars, while having limited overlaps, are complementary. The bar chart below, illustrates this by showing the complementary nature of ESG screening and controversies at MSCI. The overlap between controversial names and poor ESG names is limited, although it is clear that controversial stocks have a poorer rating on average. Each index pillar therefore should be considered separately and has its own academically backed “raison d’être” in terms of financial performance.

*FIGURE 12. ESG RATING BREAKDOWN OF MSCI WORLD AND STOCKS EXCLUDED FROM MSCI WORLD DUE TO CONTROVERSIES (SCORE <37)*

Source: DWS International GmbH, as of 30/09/2020, based on historic index holdings for MSCI Low Carbon SRI Leaders Series. Past performance, actual or simulated, is not a reliable indicator of future performance.

The view from academia: all ESG dimensions have strong academic backing

Let us step back and look at the consensus in the academic profession about the financial performance of ESG and at how a typical “dark green” index is constructed so it could harness such financial performance.

Such consensus around the link between Corporate Financial Performance (“CFP”) and ESG criteria is clear, with a seminal meta-study conducted by DWS in conjunction with the University of Hamburg in 2015 finding over 90% of over 2,000 academic studies published on the subject since 1970 showed a positive or neutral correlation between ESG factors and financial performance.

Studies around ESG ratings: Further research by BAML (2017), Goldman Sachs (2017), and multiple studies by MSCI (2019, 2020) all showed positive contributions of ESG factors, and in particular, ESG ratings to financial performance. Findings suggested that companies with strong ESG metrics outperformed their regional and sector peers, exhibiting superior risk metrics and lower future earnings volatility than those with inferior ESG scores.

Studies around controversies: The authors of the DWS/Hamburg University study further expanded their work in 2018, diving into specific ESG dimensions and isolating the correlations with CFP. The study identified that the ESG dimension with the strongest relation to CFP was in fact

---

7 The full DWS study ‘ESG and Corporate Financial Performance: Mapping the global landscape’, is available for download here: [https://institutional.deutscheam.com/content/-/media/k15090_AcademicInsights.uk_EMEA_RZ_Online_151201_Final_(2).pdf](https://institutional.deutscheam.com/content/-/media/k15090_AcademicInsights.uk_EMEA_RZ_Online_151201_Final_(2).pdf)

Sources: DWS International GmbH, Bloomberg, as of November 2020. Past performance is not a reliable indicator of future returns. Forecasts are not a reliable indicator of future returns. Forecasts are based on assumptions, estimates, views and hypothetical models or analyses, which might prove inaccurate or incorrect.
ESG reputation, one of multiple dimensions addressed within an ESG Index, through the integration of controversies screens.

Studies around activity screens: Stoxx in particular, also highlighted the importance of ESG-related activity screens, finding a clear positive contribution to returns through the exclusions of companies involved in tobacco, and thermal coal, regardless of region. Thermal coal and fossil fuels in general have become an increasingly important focus area for investors, with climate change becoming one of the most material ESG risks and investment opportunities. The focus has been further accentuated by the involvement of governments, regulators and central banks on this issue, and given the correlation between carbon intensity and MSCI “E” score of just -0.1 for MSCI ACWI Constituents, this has to be tackled through targeted rules within an index methodology.

These studies provide a solid foundation by confirming that every ESG dimension is a potential source of corporate financial performance.

The inclusion pillar: factor biases
Ratings are the most prominent selection criterion of an ESG index. Our analysis indicates that the inclusion of higher ESG rated companies does not introduce any material sector bias. This is mostly because the index construction, partly based on the MSCI ESG Leaders methodology, applies its inclusion approach at the level of a GICS Sector x Region segment. We however see a clear bias incurred by the inclusion approach on the factor (style) side. More specifically, we see more than 80% correlation between the combined factor biases caused by ESG laggard exclusions (B-CCC) and leader overweights (AA-AAA weights almost doubled). The most prominent biases caused by such rating inclusion approach is quality, with a clear trend towards more profitable, less leveraged stocks. A momentum bias is noticeable, however, it cannot be found in the ESG index due to the application of the controversies exclusions, as detailed below. The size bias is linked to the removal of laggards of very small weight in the index, indicating that small company size could be considered a catalyst to ESG risks. The effect at level of the overall index is negligible, though, especially as ESG laggards only make up a small portion (currently 5%) of the original benchmark.

When looking at the attribution effects of overweighting ESG leaders in the ESG indices, we find:

The aggregated allocation effects are positive throughout history and across regions, with a more material effect coming from AAA overweights, except in Emerging Markets. When looking at the attribution effects of ESG leaders on a 2-year horizon as per the table below, underlying drivers of the attribution could be split between style biases (20%), sector-level biases (45%) and selection effects (35%). Sector-level biases incurred at index level (IT, healthcare overweights), were explained earlier and can also be seen at the level of the Leaders ESG ratings.

In addition, the effect of removing ESG laggards from the ESG indices has proved broadly positive. This is most notable the CCC segment, but also in the B segment, with Emerging Markets standing out as a region where their exclusion has had the most positive allocation effect.

The overall ESG rating-driven allocation effects in Emerging Markets, combined with the quite consistent distribution of the positive single stock selection effects across market capitalization categories actually point to a large portion of the selection effects (and therefore of outperformance) being directly attributable to ESG ratings, this time in accordance with conventional wisdom. Figure 14 displays that especially poor ratings exclusions contributed positively to performance.

A notion support by a paper by STOXX (2018)

Sources: DWS International GmbH, Bloomberg, as of November 2020. Past performance is not a reliable indicator of future returns. Forecasts are not a reliable indicator of future returns. Forecasts are based on assumptions, estimates, views and hypothetical models or analyses, which might prove inaccurate or incorrect.
January 2021 / Investment Insights

TABLE 1. 2-YEAR ALLOCATION EFFECT BY SECTOR FOR DIFFERENT ESG RATINGS

<table>
<thead>
<tr>
<th>ESG Leader ratings</th>
<th>ESG Laggards ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>AA</td>
</tr>
<tr>
<td>Financials</td>
<td>-0.03</td>
</tr>
<tr>
<td>Health Care</td>
<td>0.1</td>
</tr>
<tr>
<td>Information Technology</td>
<td>0.85</td>
</tr>
<tr>
<td>Total</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Table: MSCI, Bloomberg PORT, DWS International GmbH. As at 30/09/2020. With only the selected sectors creating the most meaningful allocation effect for each rating category.

FIGURE 13. ALLOCATION BREAKDOWN BY RATINGS. REMOVALS IN THE LOWER RATED CATEGORIES CREATE MOST OF THE EXCESS RETURNS


About the importance of weight re-distribution

The MSCI ESG indices such as ESG Screened, Leaders or SRI re-weight the non-excluded stock proportionally to their market capitalization. This approach results in a “re-weighting” factor which acts consistently as a multiplier to market cap. For the MSCI LCSL indices, based on the Leaders methodology, the coefficient is around 2.5 to 3 on average, and could be greater the narrower / more prone to exclusions the index is. This re-weighting factor will act, somewhat counterintuitively, as an agent of dispersion in weights. In a simple example, a stock with 1% weight in the benchmark index will end up overweight by roughly 2% in an MSCI LCSL index, while a stock with 10 bps weight in the benchmark index will end up overweight by roughly 20bps. The overweight coefficient, if it reaches excessive levels, can to some degree pose a risk to financial outperformance at play in ESG indices. While outperformance is predominantly driven by exclusions, the reweighting is a powerful second round effect.

The chart below illustrates the reweighting effects in more detail. Specifically we show the current distribution of active weights of the MSCI World LCSL index versus MSCI World; a certain symmetry in the active weights between the exclusions (right tail) and the overweights (left tail) points to a healthy distribution of idiosyncratic bets within the index. In other words, the active share in this “dark green” ESG index is, given the exclusionary nature of the index, reasonably balanced.

FIGURE 14. DISTRIBUTION OF ACTIVE WEIGHTS BY UNDER/OVERWEIGHT VERSUS BENCHMARK FOR MSCI WORLD LCSL


The exclusion pillar: Controversies removal and activity screens

Although the results for the inclusion pillar indeed tally with consensual academic findings, the general attribution results from the previous section point to other material (or potentially material in future) performance drivers than factors or selection in EM. As displayed in figure 15 below, certain sectors experiencing high exclusion levels such as consumer staples, industrials and energy would typically find themselves underweight in a “dark green” index. We see from this chart that the two most important factors causing exclusions are (1) controversies and (2) activity screens, that is screens based on revenue generated from certain activities and in our example, carbon emissions, potential or effective, generated from such activities.

---

Sources: DWS International GmbH, Bloomberg, as of November 2020. Past performance is not a reliable indicator of future returns. Forecasts are not a reliable indicator of future returns. Forecasts are based on assumptions, estimates, views and hypothetical models or analyses, which might prove inaccurate or incorrect.
The MSCI LCSL indices exclude companies with MSCI controversies score below 3; these controversies are linked with norm violations such as UN Global Compact, expressed in terms of severity by assessing the depth of the violation (local or systemic) as well as its extent (impact on environment, population, social wellbeing etc.). This assessment is made at a company level but also feeds into the supply chain and impacts the assessment of client companies as well. With around a quarter of a benchmark index weight removed (for World and US, slightly lower in other regions), we indicated before that such filters incur high profile exclusions, with currently six out of the 10 MSCI World “mega caps” being removed on the back of such controversies. Our analysis on currently excluded stocks reveals no noticeable sector nor factor/style bias, except for a low size bias, as per the above, which compensates the high size bias created by the ratings inclusion pillar. It is also interesting to see that the controversies removal improves the overall rating profile of the index, albeit only by a small margin. When isolating the allocation effects arising from the removal of controversial stocks as a whole, which therefore entails an element of ESG rating bias and sector bias, the overall effects looks positive, which tends to counterbalance the largely negative effects raised in the previous section around the exclusions of mega caps.

The second most important source of exclusions in a “dark green” index is activity based exclusions. These include involvement in the following activities: tobacco, alcohol, gambling, controversial weapons, civilian firearms, adult entertainment, GMO on the social responsibility side, but also nuclear power, thermal coal, unconventional oil & gas extraction as well as CO₂ emission intensive and potential emission intensive activities on the environmental side. It is worth reiterating that MSCI LCSL indices are only sector-aware versus their non-ESG parent, as illustrated in the below figure12. Interestingly, these sector biases could be re-traced when comparing the distribution of the weights excluded due to their activity (green bars in the below chart). We find an almost perfect explanation of the sector biases encountered in the MSCI World LCSL index – this suggests that the sector biases are directly linked to the “dark green” ESG methodology. In fact we would go further, arguing that those biases are also necessary: especially given increased focus into potential emissions and fossil fuel reserves11 it could be seen that both for reasons of representativeness and of diversification12, these sector tilts could be highly desirable. It is worth noting that the activity screens do not cause major additional factor biases, with a low correlation between factor biases versus benchmark of the excluded stocks versus those biases of the MSCI World LCSL index. Finally, and maybe surprisingly, those excluded stocks have a relatively neutral effect on the overall ESG rating distribution of the final MSCI World LCSL index, partly because ESG ratings are industry specific.

10 The index methodology does not fully neutralize the effect of activity-based screens, especially for sectors which are deemed carbon-intensive, due to the construction of the index rules.
11 In March 2020 MSCI announced to improve climate risk profile and limit exclusions to stranded asset risk by incorporating climate considerations in the construction of several MSCI several indices: MSCI announces conclusions of consultation on potential enhancements to MSCI ESG indexes, 31/03/2020, https://www.msci.com/index-announcements
12 Within a sector/region segment where a substantial amount of capital has been removed due to activity screens, the pro-rata reweighting of each stock to ensure sector neutrality may create material idiosyncratic risks versus the main benchmark, which would not necessarily result in a sector-neutral ESG index having lower tracking error than the current ESG index.
ESG: More than one factor

Investing into a “dark green” ESG index is about making a difference. This does evidently translate into active share, both from a structural as well as size perspective. The darker green an investment needs to be, the higher the active share and the more relevant the question of potential unwanted biases. The table below recapitulates on the biases identified in the MSCI LCSL indices and suggests that, at this stage, all biases could be considered necessary, since they are a direct cause of the exclusionary and inclusionary approaches of such indices. Against established beliefs, we find that such “dark green” index outperformance may not only be generated via selection effects, but also and sometimes more materially via structural or style biases. We identified these biases as being a direct consequence of the ESG approach.

To this extent, one could conclude that reducing ESG to one “factor” would not be “telling the full story”\(^\text{13}\). One could also conclude that the various sources of outperformance identified provide a source of robust excess returns of such ESG indices going forward. But again, contrarily to established beliefs, selection effects on their own cannot automatically be caused by ESG ratings. They can, especially in regions like Emerging Markets where rating inclusion generates the greatest active share\(^\text{14}\). However, selection effects can also originate from mega-caps removals and equally (even pre-dominantly for more concentrated index approaches) from large over-weights in a market-cap oriented index rulebook. This risk is less intuitive and should be considered with caution. This should not be considered a critic of the predominant ESG index methodologies available, though, as market cap weighting /re-weighting has many established advantages and again, active share may be wisely spread across all different types of risk. On the contrary, investors should be confident that by investing into “dark green” indices, they could make a difference in terms of ESG metrics within a diversified and representative index framework. It is important, though, that investors remain vigilant in selecting ESG indices and ensuring that performance could be, to the best extent, directly attributed to their targeted ESG improvements.

<table>
<thead>
<tr>
<th>Main Structural bias</th>
<th>Mainly sector</th>
<th>Idiosyncratic risks</th>
<th>Performance impact</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main Style bias</td>
<td>None</td>
<td>Size bias, largely immaterial</td>
<td>Leverage, profitability, lower beta</td>
<td>Leverage, profitability</td>
</tr>
<tr>
<td>Performance impact</td>
<td>Positive</td>
<td>Mixed / Negative on large caps</td>
<td>Mixed / Positive on CCC removals</td>
<td>Mixed / Positive in AAA segment</td>
</tr>
<tr>
<td>Impact</td>
<td>Risk reduction</td>
<td>Opportunity costs &amp; Risk reduction</td>
<td>Risk reduction</td>
<td>Financial outperformance</td>
</tr>
</tbody>
</table>

**TABLE 2. ESG INDEX RULES AND EFFECTS ON PERFORMANCE**

Source: DWS International GmbH, as of November 2020

\(^{13}\) Giese, Nagy, and Lee (MSCI 2020) indicates similar conclusions, cf. p42

\(^{14}\) Bush, Chen, and Legunn (DWS 2020) infers similar findings cf. p5
Related Readings


DWS. (2020). *ESG investing - getting under the hood*.


---

The Authors

Olivier Souliac  
Passive Index Strategy & Analytics  
olivier.souliac@dws.com

Lukas Ahnert  
Passive Index Strategy & Analytics  
lukas.ahnert@dws.com

Timur Shaymardanov  
Passive Index Strategy & Analytics  
timur.shaymardanov@dws.com

Zohaib Saeed  
Passive Product Specialist  
zohaib.saeed@dws.com

---

Sources: DWS International GmbH, Bloomberg, as of November 2020. Past performance is not a reliable indicator of future returns. Forecasts are not a reliable indicator of future returns. Forecasts are based on assumptions, estimates, views and hypothetical models or analyses, which might prove inaccurate or incorrect.