SUSTAINABLE INVESTING

White Paper #3:
Big Data and Artificial Intelligence for Attractive ESG Investment Solutions

For LFIS: Arnaud Sarfati, CEO, Luc Dumontier, Senior Advisor, and Damien Vergnaud, Portfolio Manager

In collaboration with SESAMm

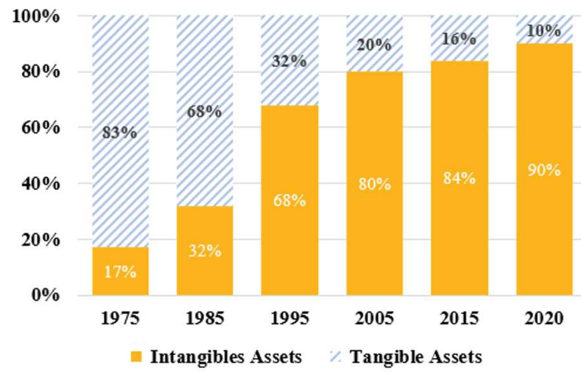
ESG news flow is all the more important as intangible assets have become predominant in the value of companies

The first paper in this series on sustainable investing[1] discussed the "zoo" of standard ESG data and its drawbacks in terms of homogeneity, comprehensiveness, opacity, point-in-time, and reactivity. This explains the contradictory conclusions reached by academic researchers and practitioners on the question of if integrating standard ESG data in investment decision making can improve risk adjusted returns.

The second paper[2] looked at this question using a factor investing framework rather than historical simulations. This analysis found that investment decisions linked to standard ESG data do not result in positive exposure to known factors (i.e., value, size, momentum, low risk) or to a new "ESG-labeled" risk factor. As such, there is no reason to believe that portfolios' returns based on standard ESG data should be positive.

However, no one disputes that ESG information can significantly impact short-term asset prices. Controversies are a good illustration to this. Among the best-known recent examples are Volkswagen (emissions scandal), Lafarge (Daesh financing), Valeant (accounting), Bayer (pesticide claims), Shell (Nigerian corruption scandal), Renault (financial wrongdoing allegations against Carlos Ghosn), etc. In each of these cases, the share price fell quickly and sharply under the threat of legal action and/or consumer boycotts. This illustrates the growing importance of intangible assets (i.e., brand, research and development, intellectual property, etc.) in corporate valuations. Exhibit 1 shows that intangible assets represented only 17% of market value for the companies in the S&P500 index in 1975. Today, intangible factors account for 90% of a company's market value.

Exhibit 1: Components of S&P500 Market Value[3]



Source: Ocean Tomo

Standard ESG data is updated, at most, a few times a year and often lags price movements. Logically, more reactive ESG data could serve as a live proxy for evolution in intangible assets and could anticipate price trends. Today we have the ability to exploit the continuous flow of textual data on the web. Positive communications from companies (e.g., gender equality measures, carbon-neutral objectives, etc.) tends to generate goodwill while negative surprises (e.g., oil spill, lawsuit, etc.) tend to generate badwill. Accessing the full scope of textual data ensures that you capture both voluntary positive corporate communications and negative news from investigative journalism and whistleblowers.

Big data and artificial intelligence for the construction of short-term ESG signals

LFIS Capital has partnered with French FinTech SESAMm to generate alternative ESG data that is more reactive as well as transparent, proprietary, homogeneous, and point-in-time. Together we have developed a quantitative platform that analyses specific ESG keywords from news articles, blogs, and social media in real time. The result is a daily ESG score specific to each stock.

The first step is "linking" or filtering over 14 billion articles from 4 million different sources and identifying those *linked* to the specific company of interest. A company's identifiers are its name, and those of its subsidiaries and board of directors, through the Knowledge Graph. For ambiguous words, Natural Language Processing ("NLP") algorithms help separate the wheat from the chaff. "Named Entity Recognition" uses a neural model to identify news specifically linked to the corporations, e.g., keeping "Apple" the proper name and not "apple" the fruit or common name. "Named entity disambiguation" takes things a step further using a word embedding approach to further refine the results, e.g., keeping "Orange" the company and not "Orange" the city which are both proper names.

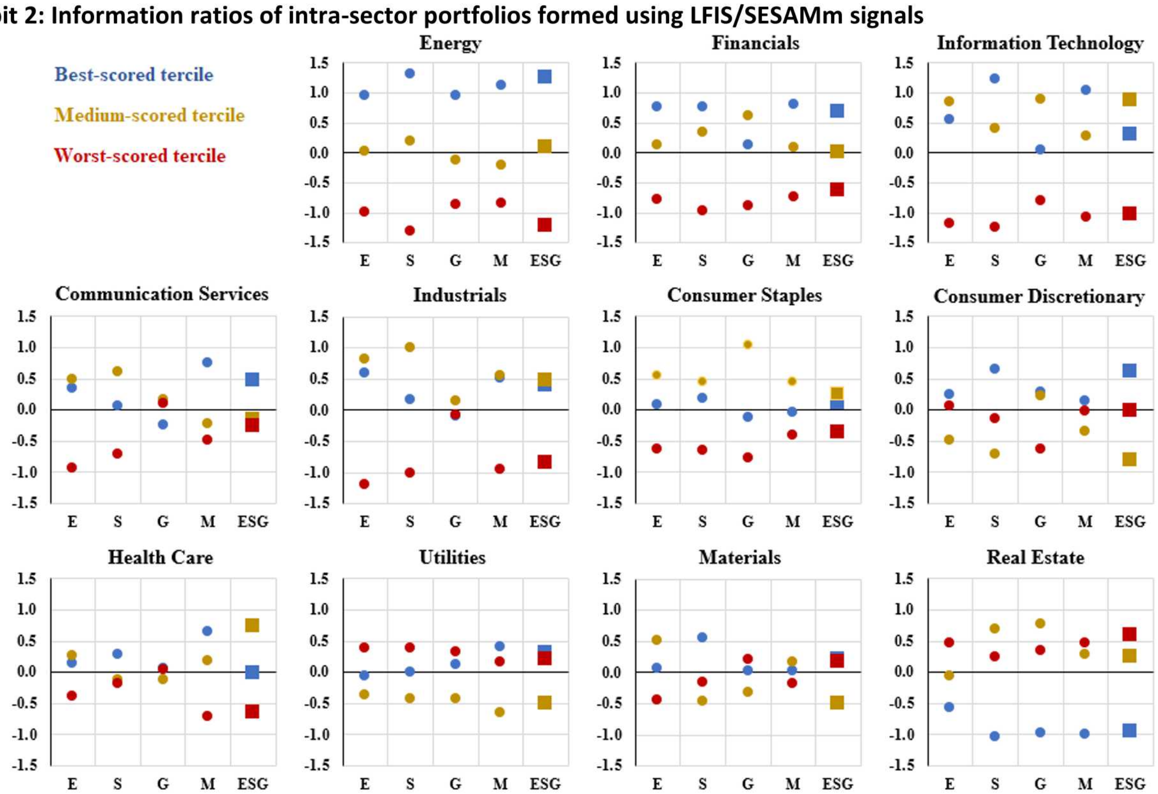
The second step is "content analysis" or screening of the articles identified in step one, based on a predefined list of ESG keywords for each thematic: E (e.g., carbon, climate change, etc.), S (e.g., diversity, gender pay ratio, etc.), G (e.g., transparency, tax evasion, etc.), and M for "miscellaneous" (e.g., scandal, reputational damage, etc.). NLP-based "sentiment analysis" attributes an indicator to each ESG mention, either negative, neutral, or positive. In addition, "ESG pertinence" quantitatively scores the relevance of each underlying article. For each day, stock, and thematic (E, S, G, and M), a score is computed by averaging the sentiment indicator of each mention with the degree of relevance of the associated article. A final signal for each thematic is calculated as exponential weighted average of these daily scores over 360 days. A final overall "ESG" signal is then calculated by averaging the thematic signals. We used an equally weighted average to avoid any overfitting bias.

Unconstrained simulations using LFIS/SESAMm ESG signals exhibit promising results

We first considered an unconstrained framework (i.e., no benchmark, daily rebalancing, etc.) to evaluate the purest potential of these signals to deliver performance. We took the stocks within the Stoxx600 index.

We started by assessing the predictive power of our signals for intra-sector allocation ("best-in-class approach"). For: (i) each day from December 2014 to December 2020, (ii) each of the 11 GICS sectors and (iii) each of the 5 thematic scores (i.e., for E, S, G, M and ESG), three equal-weighted portfolios were created. These portfolios comprised the stocks with scores in the "best", "medium", and those scoring in the "worst tercile" respectively. We then calculated the daily performance of the resulting 165 portfolios (11 sectors × 5 scores × 3 terciles) using the prior day's portfolio composition. As a benchmark, we calculated the daily performance of the 11 sectors as the equal average performance of their component stocks. A comparison of the information ratios for the 165 portfolios vs. their respective sectors over the 6-year period showed promising results for most sectors (see exhibit 2). The energy sector (top left) had the best results, with information ratios between 1 and 1.3 for the best rated portfolios (in blue). These results are uniformly higher than for the medium rated portfolios (in yellow) where information ratios range from -0.2 to 0.2. Similarly, the medium rated portfolios delivered uniformly superior results versus the lowest rated portfolio (in red), where information ratios range from -0.8 to -1.2. If only the aggregate ESG score is considered, the best-scored portfolios have significantly higher information ratios than the worst-scored in 8 of the 11 sectors. Information ratios are comparable for 2 sectors (utilities and materials), and only real estate shows a significantly lower result.

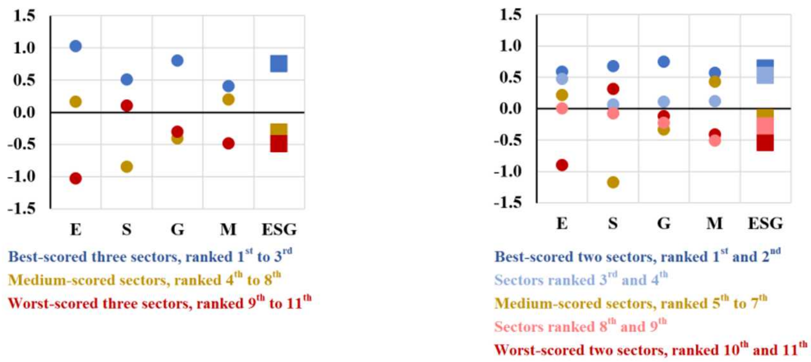
Exhibit 2: Information ratios of intra-sector portfolios formed using LFIS/SESAMm signals



Source: SESAMm, LFIS. Past performance is not an indication of future results.

We then tested the predictive power of our signals for inter-sector allocation. For each day from December 2014 to December 2020, we calculated the 5 thematic signals (i.e., E, S, G, M, and ESG) of the 11 GICS sectors as an equal average of the signals of the component stocks on the relevant date. We then applied the same methodology used for intra-sector portfolios but replacing stocks by sectors. Three equal-weighted portfolios were created: the first comprising the three best-scored sectors, the second, the next five and the third, the three worst-scored sectors. We then calculated the daily performance of the resulting 15 portfolios (5 scores × 3 clusters) using the prior day's portfolio composition. As a hypothetical benchmark, we used the daily average performance of all stocks in the investment universe. A comparison of information ratios for the 15 portfolios vs. the benchmark over the 6-year period shows promising results (see exhibit 3). For each thematic, the best-scored cluster (i.e., the three best-scored sectors) has a significantly higher information ratio than that of the worst-scored cluster (i.e., the three worst-scored sectors). For the overall ESG score, the best cluster displays a 0.6 information ratio vs. -0.5 for the worst cluster. The same study was performed considering 5 clusters instead of 3 to check the robustness of the results. The first cluster is made up of the two highest rated sectors, the second of the sectors ranked 3rd and 4th, the third of the sectors ranked 5th to 7th, the fourth of sectors ranked 8th and 9th, and the fifth of the two lowest rated sectors. The right-hand chart in exhibit 3 shows equally promising results: information ratios of the 5 clusters are ranked in the same order as their ESG signals.

Exhibit 3: Information ratios of inter-sector portfolios formed using LFIS/SESAMm scores



Source: SESAMm, LFIS. Past performance is not an indication of future results.

LFIS/SESAMm ESG signals help design attractive long only and long/short strategies

Our signals therefore help to explain the cross-section of expected stock returns within an unconstrained framework. To assess their potential as a basis for investment strategies, we considered long-only, long-short and 130/30 portfolios. The paper portfolios were rebalanced daily, based on the overall "ESG" signal for each underlying stock. Simulations were run over the period from December 2015 through December 2020 and results include transaction costs. Sector neutrality was imposed to reduce any extraneous 'noise' in the results.

The long-only ESG strategy had similar sector weights as the Stoxx600 ESG-X index. Within each of the 11 GICS sectors, we: (i) ranked stocks in descending order based on ESG signal, (ii) selected the highest-scoring stocks until reaching 20% of the sector market capitalization, and (iii) weighted the allocation to select stocks according to their relative market capitalizations. The results are in exhibit 4. This long-only strategy delivered a 7.9% annualized return, 2.9% higher than the benchmark for similar annualized volatility (17.3% vs. 17.1%). With a tracking error of 2.8%, the information ratio of the strategy is greater than 1. Results for the most recent three years are particularly convincing, reflecting growing interest and news-flow around the ESG theme.

Exhibit 4: Simulated results of long-only ESG strategy

Source: Bloomberg, LFIS, SESAMm. Past performance is not an indication of future results.

We also designed a long/short, sector neutral ESG strategy. Here, the investment universe was all stocks in the Stoxx600 ESG-X index with a capitalization of over €7.5bn. With this filter, we avoid as much as possible short squeeze phenomena, and we are quite confident that stocks could be borrowed historically. Sector weights for both the long and the short legs of the strategy are based on the number of stocks in the reference index. Within each sector, the long leg is comprised of the 20% best ESG stocks, and the short leg is made up of the 20% worst ESG stocks. Selected stocks are equally weighted for both legs. Exhibit 5 shows that the long/short investment strategy delivered a Sharpe ratio of approximatively 1 over the past 5 years, with annualized return and volatility of 6.1% and 5.9%, respectively. Like the long-only strategy, returns are very particularly robust over the past 3 years: +6.0% in 2018, +7.3% in 2019, and +11.3% in 2020.

Exhibit 5: Simulated results of long/short ESG strategy

Source: Bloomberg, LFIS, SESAMm. Past performance is not an indication of future results.

Finally, we created a "130/30" ESG strategy by simply combined 100% of the long-only ESG strategy and 30% of the long/short ESG strategy. Exhibit 6 shows that this strategy delivered a 10.8% annualized return, 5.8% higher than that of the Stoxx600 ESG-X index, with similar annualized volatility (16.9% vs. 17.1%). With a tracking error of 3.8%, the information ratio of this strategy is over 1.5, with a consistent outperformance each year.

Exhibit 6: Simulated results of "130/30" ESG strategy

Source: Bloomberg, LFIS, SESAMm. Past performance is not an indication of future results.

These results indicate that the collaboration between LFIS and SESAMm have succeeded in generating ESG signals that can capture the performance potential in short-term trends. These initial findings indicate that our alternative ESG data can serve as a foundation for attractive, timely and topical investment strategies.

[1] "ESG Implementation Challenges" by Arnaud Sarfati, Luc Dumontier and Giselle Comissiong, White Paper

[2] "Expected Return of Standard ESG Investments" by Arnaud Sarfati, Luc Dumontier and Damien Vergnaud, White paper

[3] "Intangible Asset Market Value Study" from Ocean Tomo