SUSTAINABLE INVESTING
White Paper #2:
Expected Return of Standard ESG Investments

By Arnaud Sarfati, CEO, Luc Dumontier, Senior Advisor, and Damien Vergnaud, Portfolio Manager

The causal link between standard ESG scores and financial performance is unproven.

The first white paper of this series dedicated to sustainable investing[1] illustrated the absence of consistent standards among ESG data providers. This explains why numerous simulations, by a variety of academic researchers and practitioners, are contradictory on the question of whether ESG criteria can improve the risk/return of equity portfolios. According to a recent study conducted by the NYU Stern Center for Sustainable Business[2], 33% of the 1,000+ research articles published between 2015 and 2020 show a positive relationship between ESG data and investment performance. 13% found the opposite and 54% classified their findings as "neutral/mixed".

Adding to the skepticism, a study of older research articles, published before 2015[3], showed that the number of positive results dropped by 72% when you moved from "nonportfolio-based studies" (i.e., theoretical portfolios) to "portfolio-related studies" (i.e., mutual funds, indices, and long/short portfolios). As Friede et al. argue, this gap can be partially explained by the implementation constraints and costs associated with managing real portfolios. It is likely, however, that overfitting of the theorical portfolios is also to blame.

Berg et al. raise a more serious issue in an article[4] with the evocative title of "Rewriting History II: The (Un)predictable Past of ESG Ratings". Using data sets from 2018 and 2020, this study reveals that major ESG data providers systematically restate their ESG ratings based on past performance. Their analysis explains that "score changes might have been data-mined such that firms that performed better in a given year experienced an ex-post upgrade in their […] score once the data was rewritten." This is similar to the issues portfolio managers faced a few years ago with ex-post adjustments to the earnings estimates in major databases.

Even if all studies were to agree on the correlation between ESG data and investment performance, the underlying studies have two other major limitations. First, the use of very low frequency data1 (i.e., ratings are updated few times a year, on average) and a historical observation period limited to the last decade. Such a low level of portfolio rebalancing does not allow for robust statistical conclusions[5]. Secondly, strong inflows into ESG investment solutions in recent years have mechanically fueled performance. This period may not, therefore, be representative of the more mature stage ahead for the industry, with more moderate inflows into ESG solutions. This raises the issue that prior research failed to differentiate between correlation and causality. Positive correlation between ESG indicators and investment performance has generally been interpreted as ESG (the cause) leading to investment performance (the effect). Yet, the inverse could also be true. Higher valuations could indicate successful companies with more money to invest in sustainability-related areas, leading to higher ESG ratings.

Taking all of this into consideration, we have chosen not to present an umpteenth simulation which would be just as questionable as the previous ones. Instead, this second paper presents an alternative path to assess the expected contribution of standard ESG data: factor investing.

A factor investing framework for assessing the risk/return potential of traditional ESG data.

We know for many years that the few equity portfolio managers who succeed to beat their benchmark are often those who bias their investments towards alternative factors, such as value, size, momentum, and low risk. These factors are remunerated over the long term to carry an additional risk factor that cannot be diversified away (so-called "risk premia") or to benefit from biases linked to market participants' behavior, investment constraints and structural flows (so-called "style premia").

Among the most recent academic studies, Bender et al.[6] showed that a handful of alternative premia indexes accounted for as much as 80% of "alpha" of US equity mutual funds from 2002 to 2012. This finding repeats in long-short portfolios. In Harvey et al.[7], we find the performance of equity hedge funds from 1996 to 2014, whether systematic or discretionary, was almost only attributable to their exposure to a standard set of traditional factors.

Accordingly, investment decisions linked to standard ESG data would be profitable as soon as they induce either positive exposure to factors already discovered (i.e., value, size, momentum, low risk) or exposure to a new remunerated "ESG-labeled" factor. We will successively try to validate or refute these two hypotheses by relying as much as possible on statistical analyzes.

ESG investments through the factor lens

Given that the performance of any investment can be assessed through the prism of its exposure to alternative factors, we can simply analyze the ESG signature of these factors. If alternative factors exhibit positive, structural biases to ESG, proponents of factor investing could claim to have been integrating ESG considerations for decades without even knowing it, exactly as Mister Jourdain "[has] been speaking prose while knowing nothing of it".

"My faith! For more than forty years I have been speaking prose while knowing nothing of it, and I am the most obliged person in the world to you for telling me so."
Mister Jourdain in "Le Bourgeois Gentilhomme" by Molière[8]

To look at ESG investments through the factor lens, we applied a three-step procedure. First, data was gathered on all stocks in two representative indices (the S&P500 and Stoxx600) and from well-known ESG data providers (Sustainalytics and Arabesque) at the end of each semester over the last 8 years. We then calculated 4 "point-in-time" factorial scores for each stock, considering an aggregation of price-earnings and price-to-book ratios for the value factor, market capitalization for the size factor, performance over the last 12 months skipping the most recent one for the momentum factor, and weekly volatility over the last year for the low-risk factor. Each stock was assigned a rank between -1 and

+1 for each of its individual factorial scores. Finally, we built groups of 30 homogeneous stocks based on the factorial scores to smooth results. For example, the 30 smallest capitalization stocks and the 30 largest capitalization stocks make up the first and last groups respectively of the size factor. Likewise, the 30 most volatile stocks make up the last group of the low-risk factor, with stocks with volatilities ranked from 2 to 31 making up the penultimate group, etc. For each of these groups, we calculated the equal-weighted average of the ESG scores ranging from 0 to 100. Exhibit 1 shows the results for the stocks of the S&P500 index at the end of 2020 using Sustainalytics ESG database.

Exhibit 1: ESG score (Sustainalytics) vs. factorial score for stocks in the S&P500 index at the end of 2020
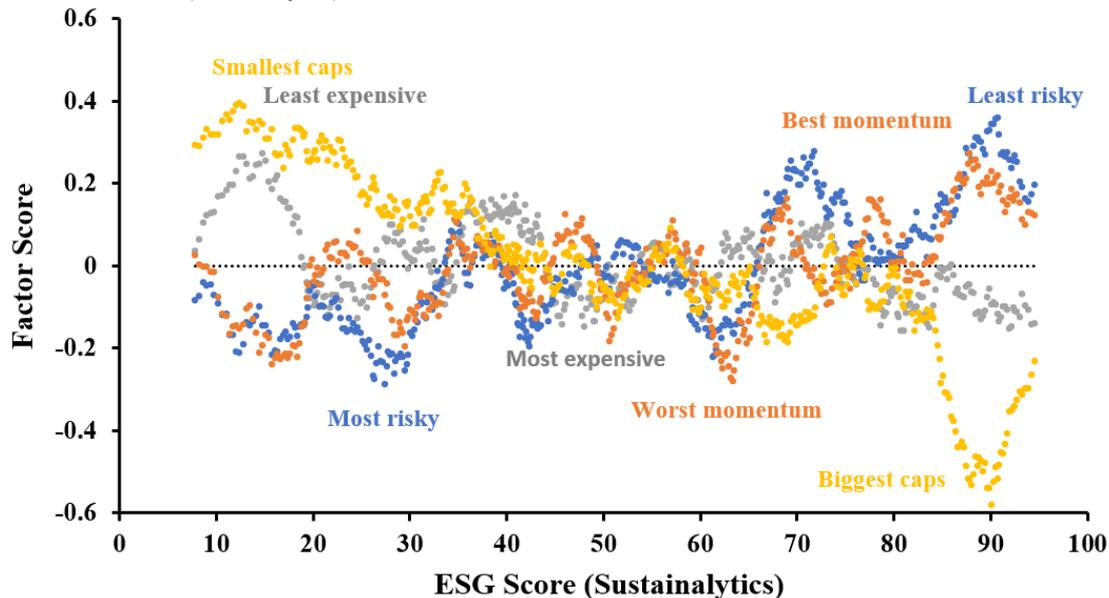

The results for the size factor, and, to a lesser extent the low-risk factor are immediately interesting. Looking at these results, the lower the capitalization, the lower the ESG score. The size factor – long the smallest caps and short the largest – is therefore strongly and negatively exposed to the ESG score. The result is the opposite for the low-risk factor. Here, the less volatile the stocks, the higher their ESG score. The low risk

factor – long the least volatile stocks and short the most volatile stocks – therefore exhibits a positive ESG bias, though of a lower absolute intensity than for the size factor. Value (long the least expensive stocks / short the most expensive stocks) and momentum (long the best momentum stocks / short the worst momentum stocks) factors do not seem to exhibit significant ESG biases.

The lower volatility of "good" ESG stocks is often cited as one of the major advantages of sustainable investing. However, we would argue that the link between good ESG stocks and low-volatility stocks might result instead from ESG being more tilted towards large-cap stocks, which are themselves less volatile. Exhibit 2 illustrates this for stocks in the S&P500 index over 2020 using Sustainalytics ESG data provider. The average volatility of the smallest stocks is higher than the average volatility of the largest, while larger companies tend to score higher for ESG. Within each quintile of market capitalization, however, the volatility spread between low ESG and high ESG stocks is not significant. This indicates that ESG is more linked to the size rather than low-volatility.

Exhibit 2: ESG score (Sustainalytics) vs. low volatility for S&P500 stocks
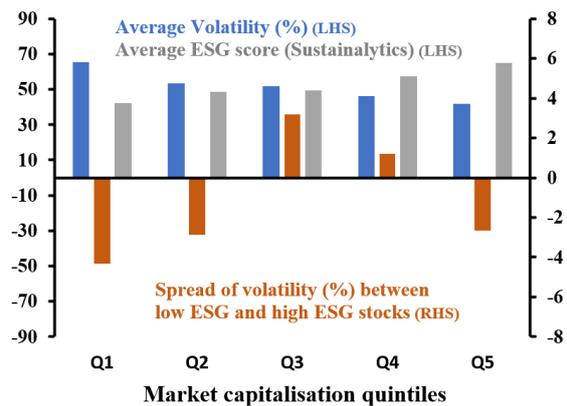
The negative exposure of ESG to the size factor can also be questioned. First, larger companies may receive better ESG scores because they can dedicate greater resources to preparing and publishing ESG disclosures. This can skew scores, rewarding large firms with higher ratings while penalizing smaller companies with limited resources. Secondly, the size bias is not as marked if data from other ESG providers is used. Exhibit 3 takes 4 factors and the S&P500 and Stoxx600 indices at the end of each semester since December 2013 and calculates the ratio between the average ESG rating of the 50

stocks with the best factorial scores and the average ESG rating of the 50 stocks with the lowest factorial score. For the size factor, this is the ratio of the average ESG rating of the 50 smallest capitalizations and the average ESG rating of the 50 highest capitalizations. The grey and yellow curves of the top left chart show a structural negative size bias in Sustainalytics ESG ratings for the S&P500 and the Stoxx600 indices, respectively. However, this bias does not exist for Arabesque ESG ratings for either index.

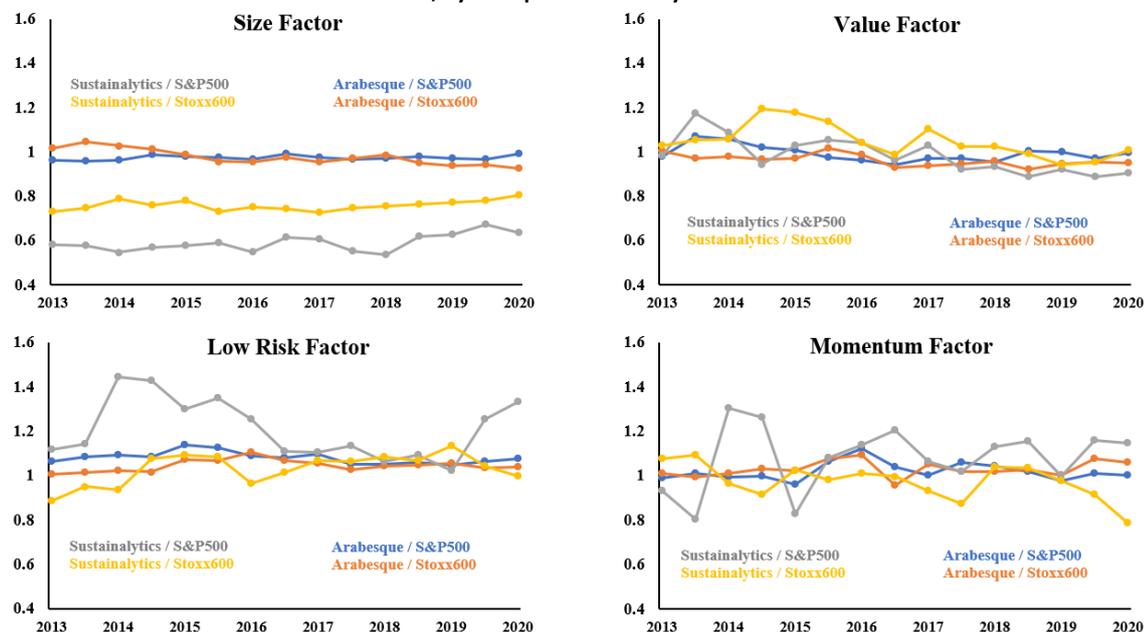**Exhibit 3: ESG scores of alternative factors, by data provider and by index**

Exhibit 3 also confirms that other alternative factors (i.e., value, low risk, and momentum) do not present any significant ESG bias, irrespective of data provider, index or date considered. This is somewhat surprising for the value factor, especially at the end of December 2020. Indeed, stocks widely considered as advanced in terms of ESG performed strongly in 2020 and became expensive in terms of valuation multiples. For example, the S&P Global Clean Energy TR index surged +141.3%

and its price-earnings ratio reached 63 at the end of 2020. However, exhibit 4 shows that the average ESG rating of stocks in the S&P Global Clean Energy index is comparable to that of stocks in both the S&P500 and S&P500 Energy indices, based on the most recent ESG data. We used Arabesque over Sustainalytics as the disclosure of data for the S&P Global Clean Energy index is higher: 72% vs. 23%.

This raises an additional flag: either certain investments that are supposed to be sustainable are not, or ESG scores do not properly reflect the sustainability of certain activities.

In the end, this analysis does not find any structural exposure to alternative factors in ESG data[9].

Exhibit 4: Average Arabesque ESG rating

	E	S	G
S&P Global Clean Energy	61	55	48
S&P500	59	57	52
S&P500 Energy	60	57	57

Assessment of the potential of standard ESG investments to represent a factor in its own right

Factor investing is a powerful analytical framework that allows you to separate the wheat from the chaff. A new investment theme is only a new factor if it satisfies strict qualification criteria, the main ones being attractive, established, orthogonal, robust, and explainable.

"Attractive" refers to the investment theme's potential to deliver positive risk-adjusted returns over the long-term. By "established" we mean that the theme is supported by a solid foundation of academic and practitioner literature. A conventional form of implementation should also exist. As previously discussed, ESG investments using traditional data do not conform to these criteria. Simulations using standard ESG data are performed over short horizons and deliver contradictory results, not least because of a wide variety of implementation approaches (e.g., best-in-class vs. best-in-universe).

"Orthogonal" refers to the fact that the investment theme should but genuinely new and not new implementation of a previously existing factor. The results above seem to validate this criterium. Traditional ESG scores do not exhibit strong structural exposure to existing factors.

"Robust" means that the investment theme is widespread and does not depend overly on implementation choices. To settle this question, we started by computing the daily excess return of ESG indices versus traditional indices by region[10] (U.S. and Europe) using MSCI as the data provider, and by provider[11] (Euronext and MSCI) in Europe, between 2013 and 2020. We then calculated the rolling 6-month correlation of excess returns between regions (i.e., U.S. vs. Europe) and providers (i.e., Euronext vs. MSCI). Exhibit 5 shows the results, specifically, very low correlation. This points to a lack of robustness.

Exhibit 5: Correlations of ESG investments between regions and providers
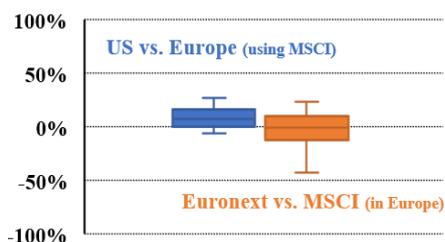
"Explainable" refers to the fact that the investment theme should either remunerate investors for exposure to an additional risk factor that cannot be diversified away ("risk premia") or stem from biases linked to market participants' behavior, investment constraints and structural flows ("style premia").

Inflows into ESG-focused vehicles lead to price appreciation for the shares of those firms with strong ESG engagement. If their earnings do not keep pace with this appreciation, valuation multiples will increase, indicating that investors should expect lower future returns over the long term. In other words, as investors pay higher prices for good ESG companies, there might exist a premium for investors who hold companies with poor ESG practices. This premium compensates investors for any ESG-related risks, such as environmental disasters or corporate scandals due to governance failures. In this case, ESG investing would be the opposite of a risk premium.

If on the contrary, earnings of "good" ESG companies improve in line with, or more, than their stock prices, valuation multiples will remain stable or improve (i.e., decrease). Giese *et al.*[12] argue that companies with strong ESG profiles are more competitive than their peers thanks to better use of resources, human capital development, and innovation management. Companies with high ESG-ratings use these competitive advantages to generate additional returns, which ultimately leads to higher profitability. In this case, ESG investing could be considered as a style factor like the price momentum factor which is often explained by a positive earnings momentum. Less controversies could also lead to reduced volatility for "good" ESG stocks. These stocks could then ultimately outperform in line with the "low risk / quality" style premium.

It is, however, difficult to consider ESG investing using traditional data as an alternative factor as it does not meet all the qualification criteria. Within the factor universe, ESG is closest to a style factor.

To conclude, the expected return of ESG investing using standard data is difficult to predict. There is no positive exposure to existing factors or to a new "ESG-labeled" factor. The next paper in this series will introduce a new data set, one developed in partnership with French FinTech firm SESAMm. This alternative data leverages on machine learning ("ML") and natural language ("NLP") processing for greater speed, transparency and reliability and is the foundation for LFIS' ESG investment proposal.

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

[2] "ESG and Financial Performance" by Tensie Whelan, Ulrich Atz, Tracy Van Holt and Casey Clark, White Paper

[3] "ESG and Financial Performance: Aggregated Evidence from more than 2000 Empirical Studies" by Gunnar Friede, Timo Busch and Alexander Bassen, Journal of Sustainable Finance & Investment 2015

[4] "Rewriting History II: The (Un)Predictable Past of ESG Ratings" by Florian Berg, Kornelia Fabisik, and Zacharias Sautner, White Paper

[5] See for example "The Sharpe Ratio Efficient Frontier" by David H. Bailey and Marcos López de Prado, Journal of Risk Winter 2012/2013

[6] "Can Alpha Be Captured by Risk Premia?" by Jennifer Bender, P. Brett Hammond and William Mok, The Journal of Portfolio Management Winter 2014

[7] "Man vs. Machine: Comparing Discretionary and Systematic Hedge Fund Performance" by Campbell R. Harvey, Sandry Rattray, Andrew Sinclair and Otto Van Hemert, The Journal of Portfolio Management Summer 2017

[8] Translated from the French: « Par ma foi ! Il y a plus de quarante ans que je dis de la prose sans que j'en susse rien, et je vous suis le plus obligé du monde de m'avoir appris cela. »

[9] This result is in line with those of Morningstar in "ESG as a Factor: Evaluating ESG Effects on Risk and Return"

[10] (MSCI USA SRI TR - MSCI USA TR), and (MSCI Europe SRI TR - MSCI Europe TR)

[11] (Euronext Euro 50 ESG EW TR - Euronext Eurozone 60 EW TR), and (MSCI Europe SRI TR - MSCI Europe TR)

[12] "Foundations of ESG Investing: How ESG Affects Equity Valuation, Risk, and Performance" by Guido Giese, Linda-Eling Lee, Dimitris Melas, Zoltán Nagy, and Laura Nishikawa, The Journal of Portfolio Management, July 2019