Methodology

Sustainable Sovereign Risk Assessment Methodology

Sovereign ESG revisited

Contents

Section 1

ESG integration into the sovereign risk analysis

The Sustainable Sovereign Risk Assessment Methodology (2SRM) is based on its predecessor, the Sovereign Risk Monitor (SRM).[1] It responds to increasing market maturity, customers' needs and the World Bank's recommendations[2] and delivers a renewed and strengthened methodology.

2SRM is a quantitative, relative and systematic approach, based on 36 indicators for 151 countries, divided into three pillars of sustainable sovereign risk assessment.

Beyond Ratings calculates a score on a quarterly basis for each indicator, starting from 1999 until the end of the latest quarter. Each of the 36 indicators is the outcome of numerous adjustments – systematic to a large extent – based on public, private and proprietary data.

All indicators are combined at (i) a risk theme level and (ii) a pillar level to obtain an aggregated score, which is derived from advanced statistical and econometric techniques discussed hereafter.

Finally, the scores are aggregated from each pillar to obtain an aggregated ESG score.

1.1 The scoring framework

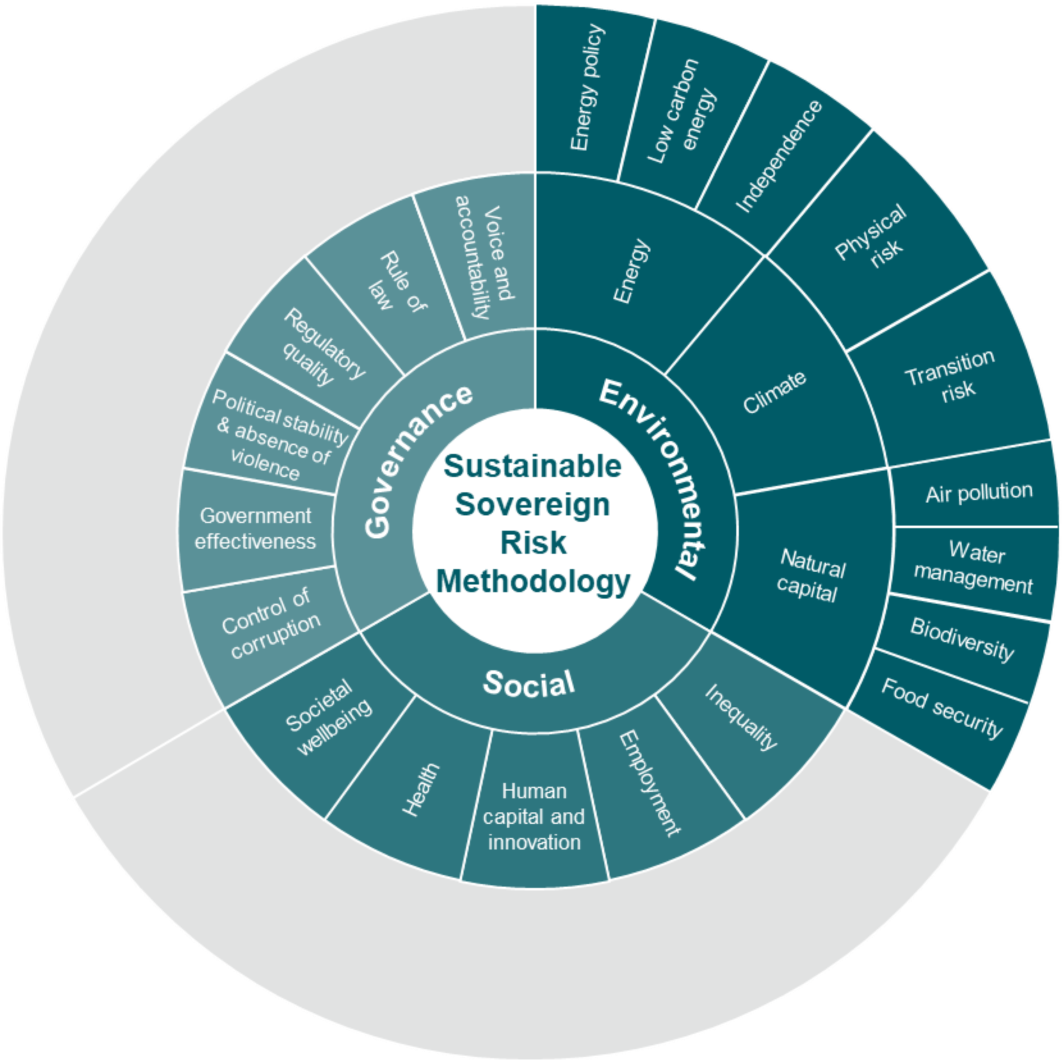
2SRM relies on the quantitative assessment of Environmental (E), Social (S) and Governance (G) pillars which characterise sovereign creditworthiness. Each pillar is structured around sub-pillars, which consist of several risk themes that include several indicators (see Figure 1).

The Environmental Pillar, for example, is represented by three sub-pillars: Energy, Climate and Natural Capital. The Energy sub-pillar is consequently made up of three risk themes: Energy Policy, Low-Carbon Energy and Energy Independence. The Energy Policy risk theme is composed of two indicators: Electricity Access and Energy Consumption. The quantity of indicators varies from one risk theme to another, but on average, they range between two and four.

[1] For more details, refer to the Sovereign Risk Monitor Methodology.

[2] Refer to the aforementioned paper, Demystifying Sovereign ESG.

Figure 1. The 2SRM framework



Source: Beyond Ratings.

1.2 The relative and systematic quantitative framework

1.2.1. From raw data to indicators

Figure 2 illustrates the general framework through which we transform raw data into indicators.

Geographical coverage

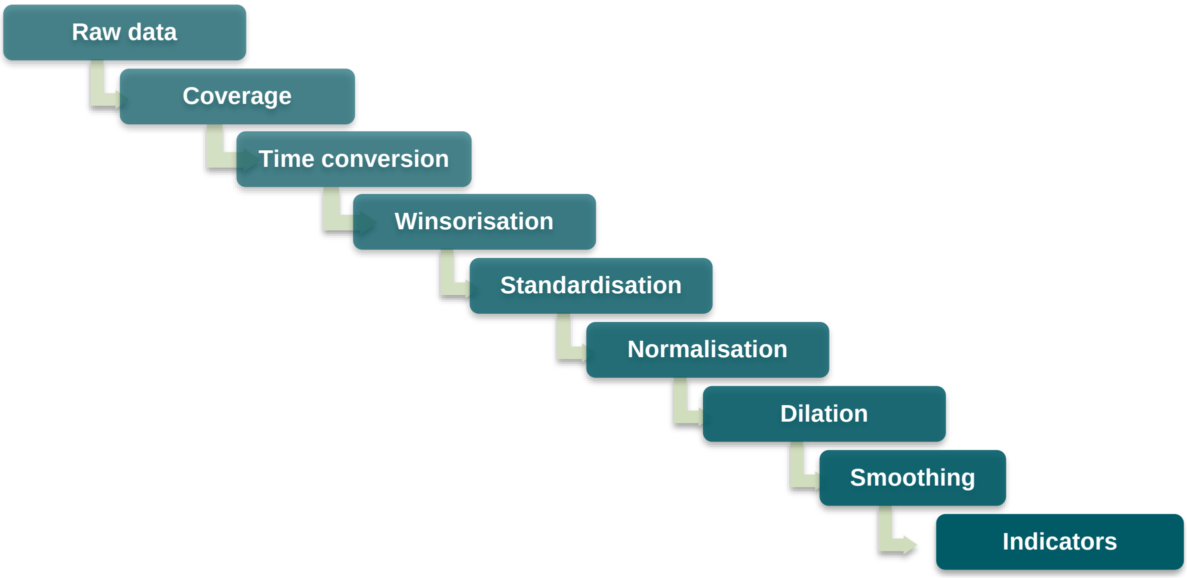
Before starting the statistical transformation of the raw data, we check that we have sufficient raw data for a given indicator and a given period for the statistical transformations to be robust enough. Thanks to a flagging process, we check for geographical coverage gaps. We first identify a flagging ratio of 25%, which is defined as the minimum threshold of geographical coverage for any indicator, for a given period. When a given indicator's coverage does not surpass this flagging ratio, it is discarded as an indicator, meaning that the raw data is then not taken into account for this period. It is important to note that this 25% flagging ratio applies to two different groups of sovereigns depending on the level of development.[3]

[3] We consider the International Monetary Fund's dynamic country classification, classifying countries as (i) Advanced Economies or (ii) Emerging Markets and Developing Economies. See section 3.2.2. and footnote 18 for more details.

Time conversion

Raw datasets are received in an annual format. As 2SRM provides scores on a quarterly basis, annual values will be attributed to the fourth quarter of a given year, which is followed by a linear interpolation between every current fourth quarter and the fourth quarter of the previous year. When the data gap is larger than the difference between two consecutive fourth quarters, the linear interpolation will only be performed for the first, second and third quarter of the given year and the fourth of the previous year. Remaining gaps will not be filled at this stage.[4]

Figure 2. From raw data to indicators



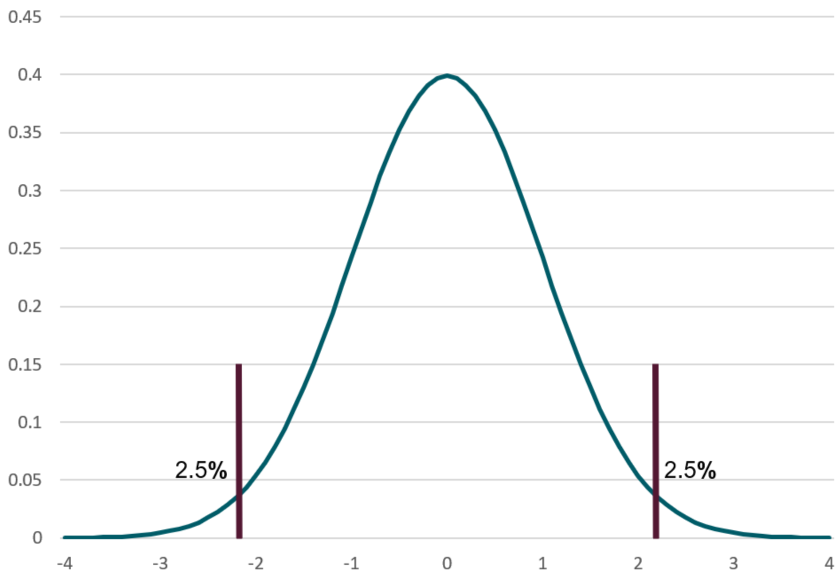
Source: Beyond Ratings.

Winsorisation

To minimise the impact of potential outliers, all indicators have been corrected for extreme values that are higher than the 97.5^{th} percentile or lower than the 2.5^{th} percentile of the distribution (see Figure 3).

[4] The Implied Temperature Rise (ITR) indicator will, however, be back backward filled for the whole time series of every country as it is a forward-looking indicator which is not recorded historically.

Figure 3. Standard distribution curve: Winsorisation process



Source: Beyond Ratings.

Standardisation

Most indicators are transformed into z-scores[5] for each country, by year and quarter. This allows us to assess the relative risk linked to the initial data and corrects for data scaling.

There are four indicators or families of indicators that are not winsorised and standardised – the World Governance Indicators produced by the World Bank, the Red List Index provided by the Sustainable Development Goals Database, the Physical Risk indicators and the Implied Temperature Rise indicator developed by Beyond Ratings – as they have already gone through a standardisation process during production.

Normalisation

The z-scores are transformed into continuous scores on an interval, ranging from 0 to 100[6] in accordance with the cumulated distribution of a standard normal distribution – where 0 represents the worst score and 100 the best score.

Two different cases provide the general framework for these additional adjustments:

(i) When the optimum is a maximum, the higher the value for the data, the higher the value of the corresponding z-score, and the higher the indicator (see Figure 4 on the left-hand side).

(ii) When the optimum is a minimum, the lower the value for the data, the lower the value of the corresponding z-score, and the higher the indicator (see Figure 4 on the right-hand side).

[5] For a raw datum denoted $X_{t,i}$ with t the date and i the country, $z\text{-}score_{X_{t,i}} = \frac{X_{t,i}-\overline{X_t}}{\sigma_{X_t}}$ with $\overline{X_t} = n^{-1}\sum_{j=1}^{n} X_t$ and $\sigma_{X_t} = \sqrt{(n-1)^{-1}\sum_{j=1}^{n}\left(X_j - \overline{X_t}\right)^2}$.

[6] The cumulated distribution of a standard normal distribution provides a value between 0 and 1 for a given z-score. This value is then multiplied by 100 for the needs of the model.

Dilation

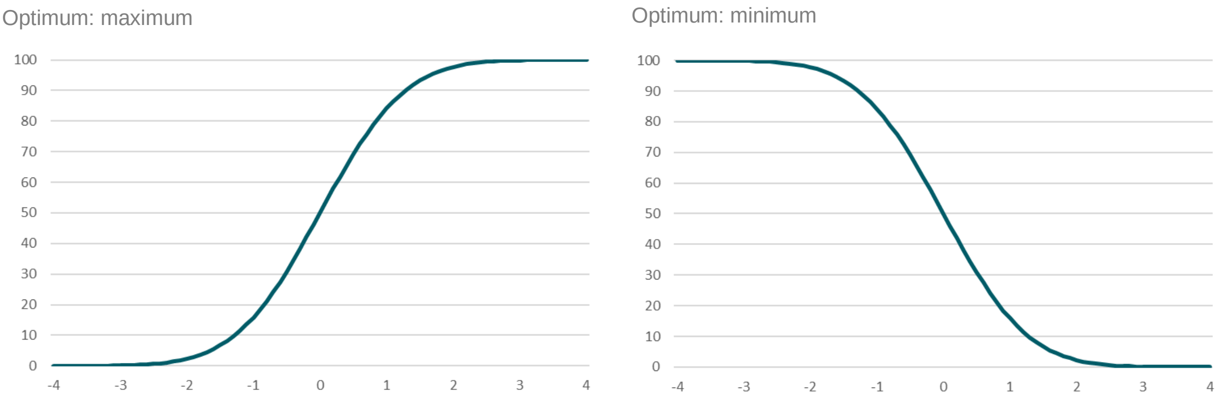
In order to maximise the discriminating power between sovereigns, a linear dilation is performed on all scores to ensure they range from 0 to 100[7]. This third phase allows us to calculate scores (*i.e.*, indicators).

The general framework detailed above does not apply at all to the Physical Risks indicators as the data already ranges from 0 to 100. However, the scale will be inverted such that higher values correspond to a higher ranking.

Smoothing

Finally, for every quarterly value for a given indicator and country, we will apply the smoothing effect. Following an exponential rule, we apply different weights to values in T, T-1, T-2 and T-3, with T being assigned the heavier weight and T-3 the lowest. This method allows us to account for fluctuations over the last four quarters and to smooth potential one-off effects or erratic data.

Figure 4. Standard Normal Cumulative Distribution Function (x axis: z-scores; y axis: scores)



Source: Beyond Ratings.

1.2.2. Aggregation – from indicators to pillars

Figure 5 illustrates the systematic approach to assigning a score to a pillar based on its underlying indicators. The chart shows how the six indicators in the Governance pillar are aggregated to provide the Governance Pillar score, representing an approach that allows us to derive a score in the form of a weighted average.

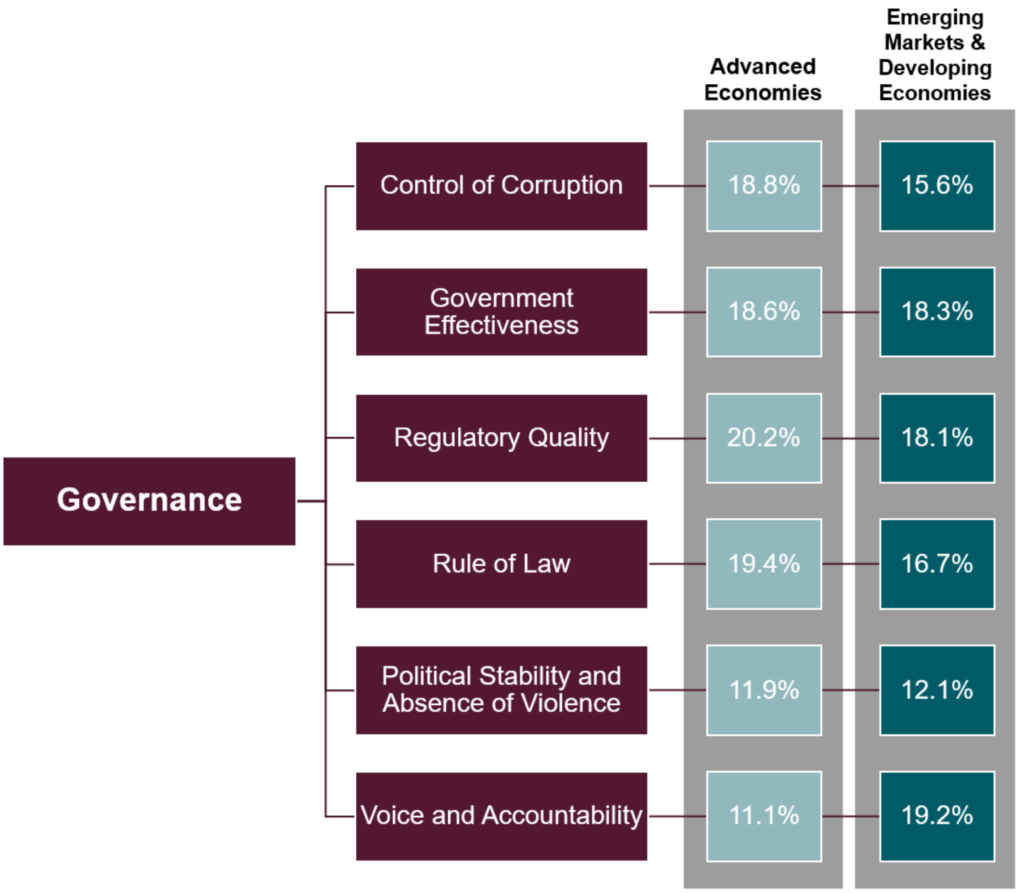
Accounting for the level of development

The relevance of each of indicator in predicting the probability of default will depend on the level of economic development for every country. As a result, we split Advanced Economies (AEs) and Emerging Markets and Developing Economies (EMDEs), as defined by the International Monetary Fund's dynamic country classification,[8] to derive an indicators' weight for each group of countries. In other words, indicators such as Poverty will be assigned a weight of 24.8% for AEs and 33.1% for EMEs.

[7] The linear dilation formula is the following: $\hat{X}_{t,i} = \frac{X_{t,i} - \min_X X_t}{\max_X X_t - \min_X X_t}$.

[8] The main criteria used by the IMF to classify the world into advanced economies and emerging market and developing economies are (i) per capita income level, (ii) export diversification (excludes resource-based economies, which hold high GDP per capita due to oil or gas rents) and (iii) degree of integration into the global financial system. This classification is updated once a year. Further information on https://www.imf.org/external/pubs/ft/weo/2020/02/weodata/groups.htm.

Figure 5. From indicators to pillars, example



Source: Beyond Ratings.

The calibration process

The weights of each indicator for each pillar, *i.e.*, intra-pillar weights, have been calibrated using the econometric modelling framework called Partial Least Squares (PLS), with an added Variable Importance in Projection (VIP) score (see Appendix 1 for further details).

This type of econometric modelling aims to be more robust than a simple linear model of the Ordinary Least Squares (OLS) type.[9] The PLS econometric model with an added VIP score allows us to take into account potential issues linked to collinearity between each indicator and to rank the information value contained in each indicator within a pillar to estimate an aggregated measure of sovereign risk.

The aggregated sovereign risk measure[10] is the endogenous variable. It is calculated as the average of non-linear numerical adjustment of the empirical default probabilities derived from the financial credit ratings of the three main credit rating agencies.[11] This aggregate measure of sovereign risk is therefore considered to calibrate the intra- and inter-pillar[12] weights of 2SRM.

[9] The OLS econometric modelling does not take into account the potential issues linked to collinearity between each indicator. Indeed, it is obvious that some indicators are strongly correlated with others, e.g., a country's general government overall balance is *de facto* strongly correlated with the general government primary balance of this same country. Therefore, the coefficients estimated through OLS are biased.

[10] This aggregated sovereign risk measure is a good proxy of a default probability.

[11] Standard & Poor's, Moody's Investors Service and Fitch Ratings, as publicly disclosed on their websites.

[12] Intra-pillar means within each pillar (e.g., within the Social performance pillar of the *Sustainability* profile) while inter-pillar means between each pillar (e.g., between the four pillars of the *Economic and Financial* profile).

Once we have estimated the coefficients for each indicator within each pillar, we normalise the scores under a significance constraint[13] to obtain a weighting set with a 100% sum for each pillar.

The results derived from this advanced econometric framework have been calibrated on a data sample from Q4 1999 to Q4 2020.[14]

Exceptions

The weights of some indicators cannot be exclusively decided using an econometric framework, but rather through the support of multidisciplinary research. This is especially the case for Physical Risk, Transition Risk, Air Pollution, Water Stress, Biodiversity and Food Security indicators.

Even if such indicators do not show a direct theoretical link to sovereign credit risk, they quantify the degree of exposure and vulnerability of countries' populations, infrastructure and ecosystems to environmental degradation, which can help predict the severity of economic impact in the short, medium and long term.

In such cases, the indicators have been attributed in equal weights to represent their equal importance, irrespective of their level of economic development.

1.2.3. From pillars to an aggregated ESG risk score

The E, S and G pillars are equally weighted to provide the ESG risk score. The reasoning for this approach is that the use of econometric regressions to derive an objective measure of the relationship between environmental indicators and sovereign credit risk assessment will be incomplete. Empirical econometric analysis tends to reduce the importance of the Environmental pillar. As a result, it seems legitimate to relatively overweight these issues and not directly attribute the weights derived from econometric models. Indeed, such issues are becoming more important to investors and are already starting to weigh on sovereign risk in those areas of the world more exposed to physical and/or transition risks due to climate change. Moreover, overall resource depletion ought to be accounted for as a set of weak signals, which are precursors for potential second-round effects in geopolitical and economic terms.

[13] In order to not underestimate too much the weight of some indicators in the modelling, we assign a minimum value ($Minimum\ Weight = 1/2N$ with N the number of indicators constituting the pillar) below which no weight can be. If some indicators are assigned that minimum weight, all the other weights are once again normalised in order to obtain weightings set the sum of which is 100% for each pillar.

[14] The calibration period runs from Q4 1999 to Q4 2020. The choice of this period was motivated by several constraints. First, we wanted to have the most up-to-date data for some of the most lagging indicators (especially in the Environmental pillar). Second, we wanted to have enough degrees of freedom for the econometric estimates. Besides, thanks to some out-of-sample estimates, we were able to highlight the strong stability of the coefficients for the regressions over time and across countries.

Section 2

By-products of 2SRM

2.1. Income-adjusted scores

Addressing the income bias

The integration of ESG criteria in sovereign credit risk analysis results in the *Ingrained Income Bias*. High-income countries (i.e., AEs) tend to have higher ESG scores and are thus favoured in ESG assessments, whereby low- and middle-income countries (i.e., EMDEs) are disadvantaged with lower ESG scores. Social and Governance scores drive this bias due to the inherent correlation with economies' level of development, while Environmental scores are less concerned due to topic diversity and divergence in assessment frameworks.

As 2SRM distinguishes between AEs and EMDEs, it allows the model to account for some of the income bias *ex ante*. However, to correct for the persisting income bias, we use an *ex post* approach[15] to generate income-adjusted E, S and G scores.

Income-adjustment methodology

We use a simple econometric framework to construct income-adjusted sovereign ESG scores. To neutralise the information related to income bias from 2SRM's E, S and G scores, we use a univariate pooled ordinary least square (POLS) regression for 149 economies, on a quarterly basis, between Q4 1999 and Q4 2020.

Empirical linkages between GNI per capita and ESG scores

We regress the Environmental, Social and Governance scores from 2SRM on the explanatory variable that is the natural logarithm of the gross national income (GNI) per capita (at purchasing power parity in constant USD) for each economy and quarter.

	Environmental	Social	Governance
$\boldsymbol{\alpha}$	38.33*** (5.94)	-69.67*** (6.30)	-100.45*** (9.81)
$\boldsymbol{\beta}$	1.90*** (0.69)	12.71*** (0.70)	16.12*** (1.11)
Observations	12,495	12,495	12,495
R-squared	0.06	0.67	0.56

Source: Beyond Ratings.

Notes: The table presents the coefficient of the POLS $E_{it} = \alpha_E + \beta_E * LN(GNI\ per\ capita_{it}) + e_{it}$ for Environmental Risk pillar, $S_{it} = \alpha_S + \beta_S * LN(GNI\ per\ capita_{it}) + s_{it}$ for Social Risk pillar and $G_{it} = \alpha_G + \beta_G * LN(GNI\ per\ capita_{it}) + g_{it}$ for Governance Risk pillar. Standard errors are in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level of confidence, respectively.

[15] For more details, refer to the Dealing with income bias in sovereign ESG scores - Sovereign ESG revisited.

Beyond the GNI per capita

We then retrieve the residuals from these three regressions (which represent the share of the initial E, S, and G scores that is not explained by the income level) to calculate the income-adjusted sovereign E, S and G scores.

To do these calculations, we transform these residuals into z-scores for each score and economy on a quarterly basis. Then, the z-scores are transformed into continuous scores based on an interval, ranging from 0 to 100, in accordance with the cumulative distribution function of a standard normal distribution – where 0 represents the worst score and 100 the best score. Finally, we can aggregate the E, S, and G scores to calculate overall income-adjusted sovereign ESG scores.

2.2. Momentum scores

Understanding ESG trends

In order to better understand contemporary trends country by country and by pillar, we have built a simple and *ex post* approach known as Momentum scoring. These Momentum scores allow for the differentiation of countries according to the recent evolutions of their respective E, S and G scores. Indeed, while the absolute levels of the E, S and G scores provide instant information, recent developments make it possible (i) to qualify the improvements or deteriorations and (ii) to anticipate future developments.

Momentum Scores Methodology

We use a simple approach based on a three-year average annual change (3Y AAC[16]) for each score of the E, S and G pillars and for each country that we compare to a rolling distribution of these same 3Y AAC over 10 years. To do this, we calculate the 10-year moving average (10Y MA) and the 10-year moving standard deviation (10Y MSTD) of the 3Y AAC. Then, we compare the 3Y AAC to its 10Y MA ± one or two 10Y MSTD. Thus, Momentum scores can take five different values:

i. If $3Y\ AAC \geq [10Y\ MA\ (3Y\ AAC) + 2 * 10Y\ MSTD\ (3Y\ AAC)]$ then $Momentum\ score\ = 2$;
ii. Else if $3Y\ AAC \geq [10Y\ MA\ (3Y\ AAC) + 10Y\ MSTD\ (3Y\ AAC)]$ then $Momentum\ score\ = 1$;
iii. Else if $3Y\ AAC \leq [10Y\ MA\ (3Y\ AAC) - 2 * 10Y\ MSTD\ (3Y\ AAC)]$ then $Momentum\ score\ = -2$;
iv. Else if $3Y\ AAC \leq [10Y\ MA\ (3Y\ AAC) - 10Y\ MSTD\ (3Y\ AAC)]$ then $Momentum\ score\ = -1$;
v. Else $Momentum\ score\ = 0$.

Finally, given the lag with which the raw data are available, we smooth these Momentum scores by using the median over four quarters to give them more simultaneity over the recent period.

A Momentum score of 2 denotes that the E, S or G score represents a strong improvement, while a Momentum score of -2 means a strong deterioration. For a Momentum score of 1 (or -1), the improvement (or deterioration) is moderate whereas for a Momentum score of 0 there is no significant change to report.

[16] For example, the three-year average annual change for the Governance score for a country i at date t is computed as follows: $3Y\ AAC\ (G_{it}) = (G_{it} - G_{it-12})/3$. Since 2SRM scores are released quarterly, we compare the score in quarter t with the same score in quarter $t - 12$ to report on the three-year period.

Appendix: Partial Least Squares (PLS) regression and Variable Importance in Projection (VIP) score

The Sovereign Risk Monitor aims to produce a comprehensive and relevant assessment of sovereign risk. To design such a framework, we have developed a statistical and econometric methodology capable of analysing multiple indicators and extracting valuable sovereign risk-related information. Then, we outline the statistical and econometric methodology and describe the key steps leading to the estimation of our two different profiles.

Sovereign risk is often influenced by numerous indicators, covering topics as wide-ranging and as different as economic performance, public finances, social features, etc. but also exposure to climate change or the quality of governance. Some indicators that make up these topics are uncorrelated, while others show a strong correlation. Therefore, extracting precise and specific sovereign risk-related information cannot be undertaken by using simple regression techniques as the results would be biased. To circumvent this issue, we use specific regression techniques to estimate the weight of each indicator in predicting an aggregated sovereign risk measure. The model we use is as follows:

$$Y = \alpha + \sum_{j=1}^{N} \beta_j X_j + \epsilon$$

where:

- Y is the aggregated sovereign risk measure with $Y = (Y_1, \ldots, Y_n)^t$, t the number of quarters and n the number of countries;
- For $j = 1, \ldots, J$, X_j is the j-th explicative indicator X matrix and J the number of indicators;
- β_j is the j-th coefficient. As already stated, it cannot be estimated by a simple Ordinary Least Squares regression as this estimator would be biased.

These indicators can present strong correlations (*e.g.*, between Social Risk and Governance Risk indicators), hence, to consider this specificity of the selected data, we use Partial Least Squares (PLS) regressions. That econometric framework, developed by Wold[17] in the 1960s, enables the construction of predictive models in the presence of many correlated independent variables. It finds orthogonal components – thus eliminating the multicollinearity issue – of the X matrix that explain as much as possible the covariance between X and Y. Then, this breakdown of X is used in the regression to predict Y.[18] More precisely, the PLS regressions follow several steps:

(i) The PLS regressions produce a matrix W such as $T = XW$, where T is the factor score matrix and W is estimated such as to minimise collinearity and maximise the covariance between the explanatory and endogenous variables;
(ii) We estimate the matrix Q so that $Y = TQ + E$;
(iii) We estimate the matrix P so that $X = TP + E'$;
(iv) We compute $\beta = WQ$.

[17] Wold, H., 1966, "Estimation of principal components and related models by iterative least squares", in P.R. Krishnaiaah (Ed.), *Multivariate analysis*, pp.391-420.
[18] Abdi, H., 2003, "Partial Least Squares (PLS) Regression", The University of Texas at Dallas.

To estimate the T matrix, the standard algorithm for computing PLS components is used, *i.e.*, Nonlinear Iterative Partial Least Squares (NIPALS) algorithm. It uses all the matrices defined above to estimate W and then compute T.

The aim is not to predict directly Y but rather to find the optimal weights of each indicator in SRM. So, the β coefficient we find in the regressions is not used directly. Instead, the Variable Importance in Projection (VIP) score is used. It represents the summary of the importance of each indicator in finding the components of the X matrix[19] during the first step of the PLS regressions. Formally, it is the weighted sum of squares of the PLS weights (the W matrix), which considers the explained variance of each dimension. It is used to select relevant predictors according to their value. In the academic literature, the VIP score is statistically significant if it based above a given threshold ranging from 0.8 to 1.[20] However, as we do not want to exclude too many indicators, we use the VIP scores directly to calculate the weights. This approach remains relevant because VIP scores higher than 0.8 account for more than 80% of SRM indicators. The last 20% are rarely below 0.5.

[19] Palermo, G., P. Piraino, and H.-D. Zucht, 2009, "Performance of PLS regression coefficients in selecting variables for each response of a multivariate PLS for omics-type data", *Advances and Applications in Bioinformatics and Chemistry: AABC*, *2*, pp. 57–70.

[20] Chong, I.G., and C.H. Jun, 2005, "Performance of some variable selection methods when multicollinearity is present", *Chemometrics and Intelligent Laboratory Systems 78*, pp. 103–112.

For more information about our indices, please visit ftserussell.com.

About FTSE Russell

FTSE Russell is a leading global index provider creating and managing a wide range of indices, data and analytic solutions to meet client needs across asset classes, style and strategies. Covering 98% of the investable market, FTSE Russell indices offer a true picture of global markets, combined with the specialist knowledge gained from developing local benchmarks around the world.

FTSE Russell index expertise and products are used extensively by institutional and retail investors globally. For over 30 years, leading asset owners, asset managers, ETF providers and investment banks have chosen FTSE Russell indices to benchmark their investment performance and create investment funds, ETFs, structured products and index-based derivatives. FTSE Russell indices also provide clients with tools for asset allocation, investment strategy analysis and risk management.

A core set of universal principles guides FTSE Russell index design and management: a transparent rules-based methodology is informed by independent committees of leading market participants. FTSE Russell is focused on index innovation and customer partnership applying the highest industry standards and embracing the IOSCO Principles. FTSE Russell is wholly owned by London Stock Exchange Group.

For more information, visit ftserussell.com.

To learn more, visit ftserussell.com; email info@ftserussell.com; or call your regional Client Service Team office:

EMEA

+44 (0) 20 7866 1810

North America

+1 877 503 6437

Asia-Pacific

Hong Kong +852 2164 3333

Tokyo +81 3 6441 1430

Sydney +61 (0) 2 8823 3521