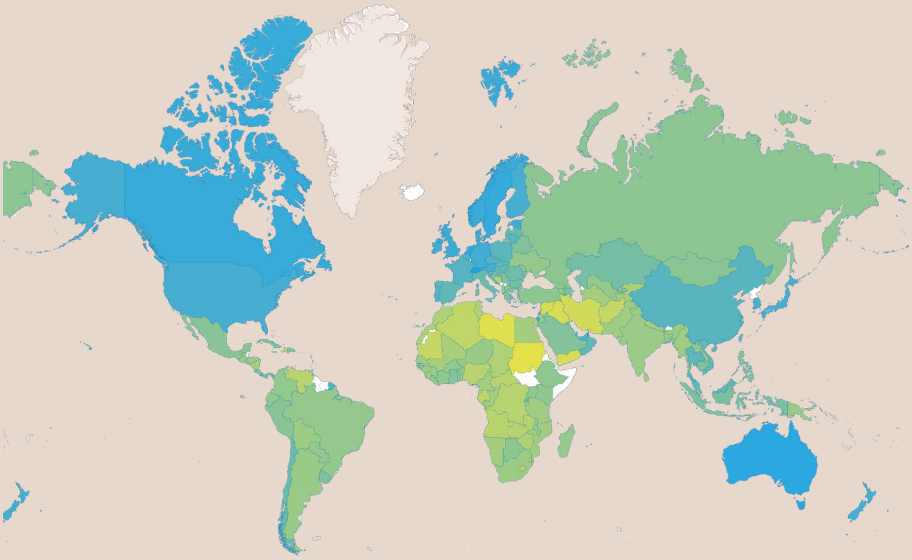
ELITE QUALITY INDEX
EQx2022

This PDF includes a contribution from the following book:

ELITE QUALITY REPORT 2022

Statistical evaluation of the Elite Quality Index

Céline Diebold, Doctoral Researcher, Institute of Economics (FGN-HSG),
University of St.Gallen, Switzerland



Cite as:

Diebold, C. (2022b). Statistical evaluation of the Elite Quality Index. In T. Casas-Klett & G. Cozzi. (Eds.), *Elite Quality Report 2022: Country Scores and Global Rankings*. Zurich: Seismo. doi: 10.33058/seismo.30769.1490

Published by Seismo Press AG, Zurich and Geneva in partnership with the Foundation for Value Creation.

©2022 the Foundation for Value Creation, St.Gallen, Switzerland, distributed under the terms and conditions of the Creative Commons license CC BY-NC-ND (http:/creativecommons.org/licenses/by-nc-nd/4.0/).

2.5 Statistical evaluation of the Elite Quality Index

"global uncertainty and sensitivity analyses are often not done.
Anyone turning to a model for insight
should demand that such analyses be conducted,
and their results be described adequately and made accessible"

(Saltelli et al., in a comment published in *Nature*, 2020, p.483)

Introduction

When the EQx was first introduced in 2020, the index provided, for the very first time, a framework to comparatively measure whether a country's elites run business models that primarily focus on value creation or value extraction. Although the first iteration of the EQx covered only 32 countries, geographical coverage surged to a remarkable 151 countries in the EQx2021, a number that has been maintained for the EQx2022. The increase in country coverage in 2021 was accompanied by an even larger increase in global media and press attention for the index. Newspaper articles appeared worldwide: in Angola,W Argentina, Brazil, China, Finland, Germany, India, Russia, Singapore, Switzerland, Turkey, Ukraine, the United States, and Kazakhstan, to name only a few. In their coverage of the EQx, journalists have generally focused their analyses on their countries' own position in the global EQx ranking. This ranking emerges as an output of the underlying model, which constructs the index.

However, when modelling any such index, uncertainties arise at each step of the construction process. For instance, how are the Indicators selected and used; how are the normalisation, weighting, and aggregation schemes chosen? This raises two key questions: how reliable and meaningful are the EQx rankings? And how do the uncertainties that underlie the index assumptions propagate to the country rankings? These questions have been addressed in a peer-reviewed paper recently published in *Social Indicator Research* that critically reflects on the EQx2021 rankings (Diebold, 2022). This article provides a summary of the main insights from that paper.

Literature and Methodology

The paper conducts a global Uncertainty and Sensitivity Analysis (UA and SA). The UA aims to quantify the overall uncertainty i.e., variation in the model output as a result of the uncertainties in the inputs, while the SA "is the study of how the uncertainty in the output of a model (…) can be apportioned to different sources of uncertainty in the model input" (Saltelli, 2002, p. 579). Global UA and SA take account of the entire variation in the range of inputs and explore all sources of uncertainty simultaneously in order to capture possible synergy effects among uncertain input factors (Iooss & Saltelli, 2015, p. 2; OECD, 2008, p. 118).

In particular, the OECD (2008, p. 117) suggests following the approach presented by Saisana et al. (2005) to conduct a global UA and SA based on a single Monte Carlo experiment, where the index is calculated numerous times based on randomly selected input factors. The input factors trigger which methodological choices are adopted. Table 1 lists the input factors, along with their associated distribution and explanation. Input factors are sampled randomly and independently, using Sobol's quasi-random sequence, following the process developed by Sobol and Kucherenko (2005).

For each set of input factors, the model is evaluated, yielding a probability distribution function (pdf) of the model output. The descriptive analysis of this pdf constitutes the UA, while the SA apportions the uncertainty in the model output to individual or the interaction of methodological choices in the index set-up.

Table 1: Summary of the uncertain input factors and their distribution

Input factor	*Definition*	*Distribution*
X_1	Imputation of missing values	$\mathcal{DU}[1,2]$ where 1 ≡ handling of missing values according to EQx2021 methodology, 2 ≡ missing values fully imputed using predictive mean matching.
X_2	Measurement error	$\mathcal{DU}[1,2]$ where 1 ≡ original raw data is used, 2 ≡ normally distributed random error with mean 0 and standard deviation (std. dev.) equal to 1/5th of an indicator's observed std. dev. is added to raw data.
X_3	Omission of individual indicators	$\mathcal{U}[0,1]$ where $[0, \frac{1}{(Q+1)})$ ≡ no indicator excluded, $[\frac{1}{(Q+1)}, \frac{2}{(Q+1)})$ ≡ exclude indicator I_1, (…) $[\frac{Q}{(Q+1)}, 1]$ ≡ exclude indicator I_Q.
X_4	Choice of conceptual optima	$\mathcal{DU}[1,2,3]$ where 1 ≡ EQx conceptual optima where applicable, 2 ≡ EQx optima plus one std. dev. of the indicator's raw data, 3 ≡ EQx optima minus one std. dev. of the indicator's raw data.
X_5	Normalisation scheme	$\mathcal{DU}[1,2]$ where (prior to rescaling, values to range from 0 to 100) 1 ≡ EQx normalisation, 2 ≡ MinMax.
X_6	Weighting scheme	$\mathcal{DU}[1,2,\ldots 9]$ where 1 ≡ EQx weighting, 2 ≡ equal indicator weights (Altern. 1), 3 ≡ equal pillar weights (Altern. 2), 4 ≡ equal indicator and pillar weights (Altern. 3), 5 ≡ equal indicator, pillar and index area weights (Altern. 4), 6 ≡ equal indicator, pillar, index area and sub-index weights (Altern. 5), 7 ≡ equal final indicator weights (Altern. 6), 8 ≡ Sub-index Power & Value at 0.25 & 0.75 (Altern. 7), 9 ≡ Sub-index Power & Value at 0.5 & 0.5 (Altern. 8), 10 ≡ equal index area and sub-index weights (Altern. 9).
X_7	Aggregation scheme	$\mathcal{DU}[1,2]$ where 1 ≡ EQx aggregation (linear), 2 ≡ geometric aggregation at sub-index level.

Note: $\mathcal{U}$ stands for uniform, and $\mathcal{DU}$ indicates a discrete uniform distribution.

Results

A total of 80,000 model runs were performed in the course of the Monte Carlo experiment. This implies that index scores and ranks are computed for all 151 countries considered by the EQx2021, for each of the 80,000 random combinations of methodological choices. Hereinafter, the index as computed using any realisation of uncertain input factors is denoted as Monte Carlo EQx (MCEQx).

Uncertainty Analysis

The pdf of the MCEQx provides the basis for the UA, where the output distribution is presented and characterised (Saisana et al., 2005, p. 310; Saltelli et al., 2019, p. 30), quantifying the uncertainty in the model output.

Figure 1 compares the EQx2021 ranking with the distribution of ranks implied by the MCEQx. The figure shows that countries ranking in the top 50 of the EQx2021 have a comparatively low variation in comparison to the MCEQx ranks; their rank seems largely robust to alternative methodological choices. However, starting with Cuba at rank # 53, the overall stability in the rankings is dramatically reduced. Countries, especially those in the middle and lower performing sections of the EQx2021 rankings exhibit a large range of possible ranks depending on the methodological set-up, implying that any interpretation or conclusion on the Elite Quality of these countries should be treated with caution. Interestingly, countries in the bottom 10 of the EQx2021 ranking again fluctuate much less when ranked using alternative methodological choices.

Sensitivity Analysis

In the next step, a variance-based SA was conducted to identify the input factors that contribute most to the variation of the model output. The paper computes Sobol' indices (Sobol', 1990), applying an estimator recently presented by Azzini et al. (2020). Figure 2 plots, for each input factor, the Sobol' sensitivity indices of first-order effect (S_i, red) and total effect (S_{iT}, blue) for the average shift in country rankings ($\overline{R}$). $\overline{R}$ is calculated as the average of the absolute difference in countries' rankings in comparison to the baseline EQx2021 ranking, for all countries. Note that to distinguish relevant input factors from less important ones, Sobol' sensitivity indices do not rely on an absolute threshold, but rather indicate the relative importance of input factors.

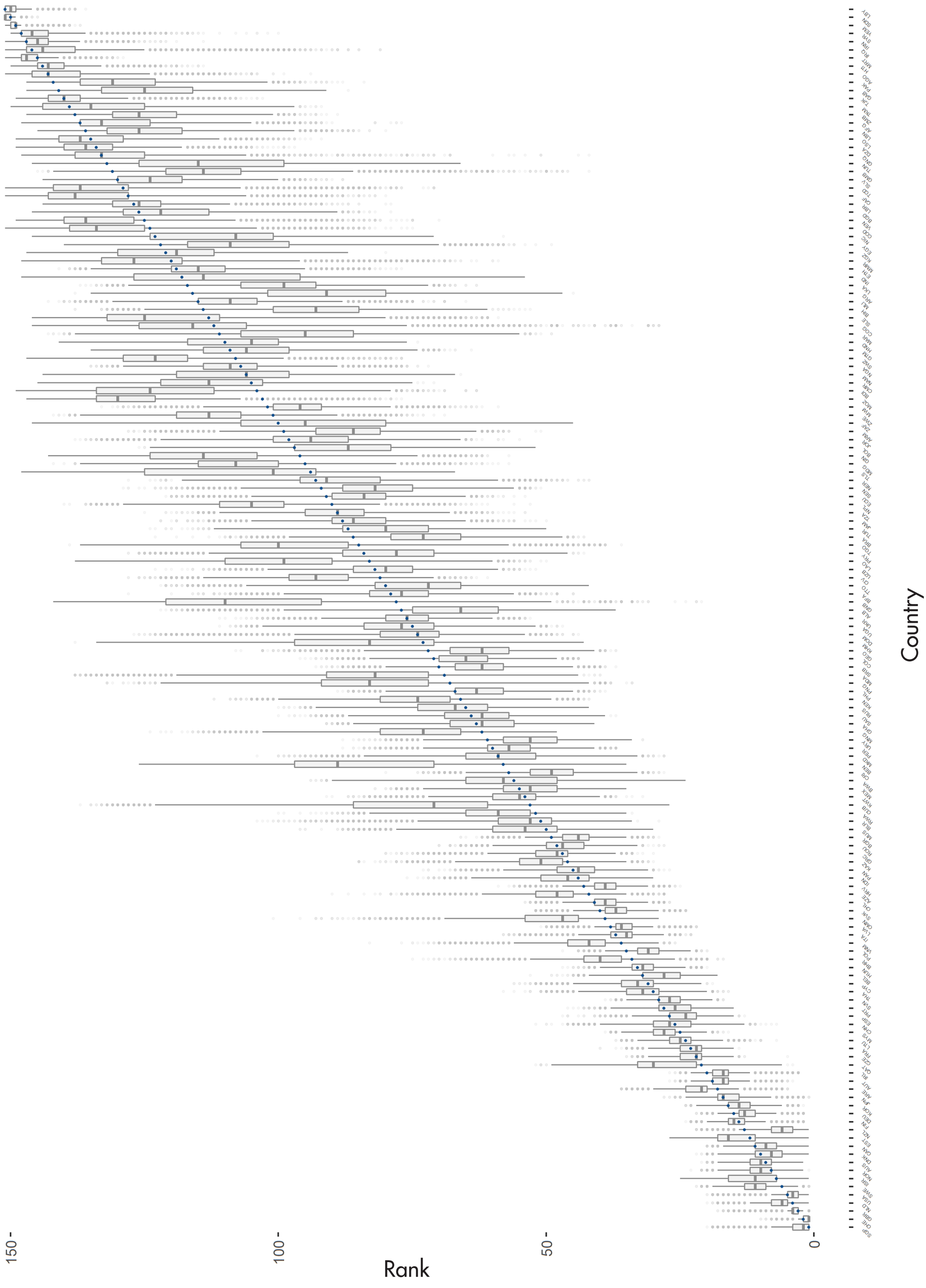
The first-order effect indicates the share by which output variance could be reduced if the considered input factor could be fixed individually. None of the first-order sensitivity indices is equal to 1, which implies that no input factor uniquely determines the index ranking. Taken individually, the input factors determining the normalisation and the weighting scheme are by far the most important, explaining roughly 20% and 40% of output variance respectively.

Next, total order effect was inspected, which adds to an input factor's S_i all interaction effects that involve said input factor. The S_{iT} reveals that the influence of the methodological choices surrounding the measurement error, the omission of indicators, the choice of conceptual optima and the aggregation scheme are relatively negligible. Accounting for interaction effects, the most influential input factors are those that trigger the imputation of missing values, the normalisation and the weighting scheme.

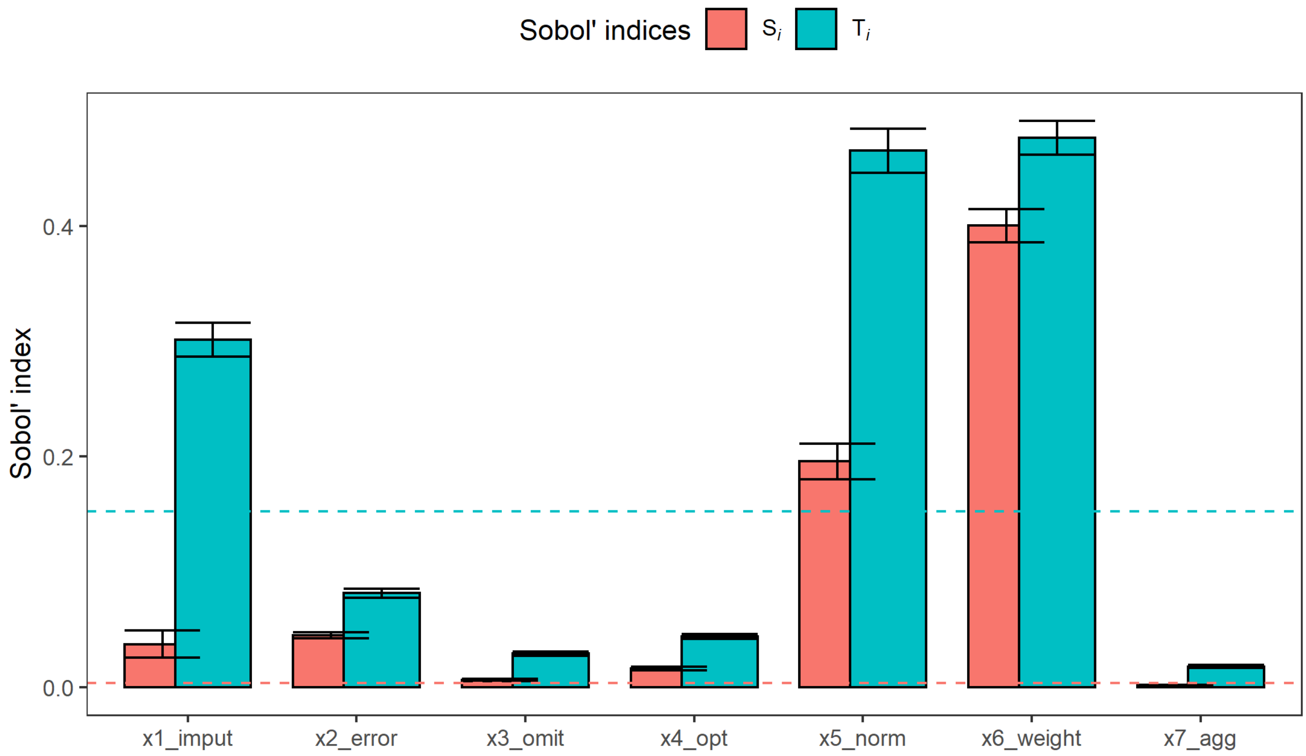
A notable difference between S_i and S_{iT} indicates interaction effects with one or several other inputs. Interestingly, the weighting scheme contributes to the output variance mainly individually, not via interaction effects, while conversely the importance of the normalisation scheme is largely due to interaction effects caused by whether missing values are imputed.

Visual 2.7: Boxplot of MCEQx ranks per country

Note: Countries are ordered by their EQx2021 rank, indicated by the blue dot. Grey boxplots illustrate the distribution of the MCEQx: boxes include 50% of a country's MCEQx ranks, and the horizontal line within a box indicates a country's median MCEQx ranking. Whiskers span up to +/- 1.5 times the interquartile range (IQR: Q3 – Q1), up to the lower observed point from the MCEQx ranks that falls within this distance. All other observed ranks are plotted as outliers (grey dots).



Visual 2.8: Sobol' sensitivity indices



Note: Whiskers denote confidence intervals obtained from the bootstrap technique. In order to estimate the numerical approximation error, Sobol' sensitivity indices for a dummy input factor that has no influence on the index ranking are also computed. The estimate of the dummy input factor is visualised through the use of a dashed line. This allows for the identification and visualization of those input factors whose contribution to the output variance is less than the approximation error.

Discussion of the Results in the Context of the EQx2022

What do the paper's results imply for the EQx methodology? Generally, it must be noted that judgement calls are inevitable throughout any index construction process. Thus, it is often difficult to argue why one scheme or methodological approach is chosen over another.

Among the three most important input factors, the weighting scheme can arguably be most convincingly rooted in the theoretical framework. Indeed, in the case of the EQx, the weightings, particularly at the index area and sub-index levels, are deduced from conceptual deliberations. On the other hand, it may be more difficult to legitimise applying one approach to missing values, or one normalisation method over another with convincing arguments. Since a large amount of uncertainty stems from the interaction between how missing values are handled and the normalisation scheme, a promising approach in improving the meaningfulness and robustness of the EQx ranking would be to reduce the index' sensitivity towards how missing values are addressed. This can most intuitively be achieved by reducing the amount of missing values in the EQx Indicator data.

Roughly 26% of all Indicator data was missing in the EQx2021. An improved procedure to impute missing values with past data has led to a reduction of this statistic by roughly 4 percentage points, to 22.4%, in the EQx2022. This has been the most important adjustment to improve the robustness and meaningfulness of the EQx2022 rankings. Nevertheless, there is a trade-off between the number of Elite Quality aspects and countries covered by the index on the one hand, and data availability and ranking robustness on the other hand. A complete reduction of the former to the detriment of the latter does not appear to be desirable, in the view of the index authors. Hence, the EQx2022 once again follows the OECD guidance (2008, p.17) in that "transparency must be the guiding principle of the entire exercise" (OECD, 2008, p. 17). The EQx2022 presents each country's performance together with the available number of datapoints that a country's EQx score is calculated from. This transparently illustrates that each index score relies on a different set of Indicators, which affects the cross-country comparability of index scores. Any interpretation or conclusion on the Elite Quality of countries with low data availability should be treated with caution. More generally, this approach allows for the interpretation of a country's EQx performance and relative position in the country rankings in the context of data availability.

Céline Diebold, Doctoral Researcher, Institute of Economics (FGN-HSG), University of St.Gallen, Switzerland