European Commission

JRC TECHNICAL REPORTS

The JRC Statistical Audit of the Social Progress Index (SPI)

Norlén, Hedvig
Caperna, Giulio

2018



Joint Research Centre

EUR 29576 EN

Contact information

Competence Centre on Composite Indicators and Scoreboards
jrc-coin@ec.europa.eu

JRC Science Hub
https://ec.europa.eu/jrc

JRC113130

EUR 29576 EN

PDF ISBN 978-92-79-98411-2 ISSN 1831-9424 doi:10.2760/343522

Luxembourg: Publications Office of the European Union, 2018

How to cite this report: Norlén Hedvig, Caperna Giulio, *The JRC Statistical Audit of the Social Progress Index (SPI)*, Publications Office of the European Union, Luxembourg, 2018, ISBN 978-92-79-98411-2, doi:10.2760/343522, JRC 113130

Contents

Abstract

A spur of social progress is enabling people to fulfil their own potential and in doing so the capability of the society they are a part of. The Social Progress Index (SPI) is an international monitoring framework for measuring social progress without resorting to the use of economic indicators. It provides a basis to understand the relationship between economic and social progress and measures country performance on aspects of social and environmental performance. The Social Progress Index builds on three dimensions: Basic Human Needs, Foundations of Wellbeing and Opportunity. These dimensions establish the basis of the framework and are used to aggregate 51 social outcome indicators organized in 12 components into a single summary measure. The statistical audit discussed in this report was conducted by the European Commission's Joint Research Centre, and it aims at maximizing the reliability and transparency of the Social Progress Index. The audit focuses on the statistical coherence and the impact of key modelling assumptions used in the SPI framework. The statistical audit of the SPI should enable policy analysts and researchers alike to draw more relevant and well-targeted conclusions regarding inclusive growth strategies that benefit everyone at all levels of economic development.

1 Introduction

The Social Progress Index (SPI) measures country performance on aspects of social and environmental performance, which are relevant for countries at all levels of economic development. The index ranks 146 countries, from less developed to well-developed, on social progress. Another 90 countries are partially ranked, and in total the index measures some aspects of social progress in almost all the world's population. Social progress is measured taking into account the following three broad aspects:

1. Meeting everyone's basic needs for food, clean water, shelter and security.
2. Living long, healthy lives with basic knowledge and communication and a clean environment.
3. Practicing equal rights and freedoms and pursuing higher education.

These three broad aspects are the main ingredients, the so called dimensions, in the index and are entitled Basic Human Needs, Foundations of Wellbeing and Opportunity. The dimensions are used to organize and aggregate 12 components and 51 social outcome indicators into a single summary measure. The SPI framework does not include economic indicators and it provides a basis to understand the relationship between economic and social progress. Traditional measures of national income, such as GDP per capita, do not capture the full picture of social progress.

The index was developed by the nonprofit organization the Social Progress Imperative based in the USA and was first released as a beta version in 2013. In 2014 the official version of the index was published and it is updated annually. The statistical audit of the 5th version of the SPI was performed by the European Commission's Competence Centre on Composite Indicators and Scoreboards[1] (COIN) at the Joint Research Centre (JRC) and was conducted upon invitation of the index developers.

The analysis herein aims at clarifying the transparency and reliability of the SPI framework and thus to enable policymakers to derive more accurate and meaningful conclusions, and to potentially guide their choices on priority setting and policy formulation when it comes to shaping inclusive growth strategies. An earlier version of the SPI 2018 was evaluated in May 2018 by the JRC.

The JRC assessment[2] of the SPI 2018 focuses on two main issues: the statistical coherence of the hierarchical structure of indicators and the impact of key modelling assumptions on the SPI ranking. The JRC analysis complements the reported country rankings for SPI with intervals in order to better evaluate the robustness of these ranks to the computation methodology (in particular estimation of weights and aggregation formula at the dimension level).

[1] Advice and recommendations given by the whole COIN team and in particular by Michaela Saisana, have been included in the audit report.

[2] The JRC statistical audit is based on the recommendations of the OECD & JRC (2008) Handbook on Composite Indicators, and on more recent research from the JRC. Generally, JRC audits of composite indicators and scoreboards are conducted upon request of their developers, see https://ec.europa.eu/jrc/en/coin and https://composite-indicators.jrc.ec.europa.eu/

2 Data analysis and rationale for the choices supporting the SPI construction

Relevance to the SPI framework. The 2018 version of SPI ranks 146 countries of all levels of economic development. Another 90 countries are partially ranked however the statistical audit is concentrated on the 146 fully ranked countries. 51 social outcome indicators have been selected by the SPI developers for their relevance to a specific dimension, capturing one of three broad elements of social progress, Basic Human Needs, Foundations of Wellbeing and Opportunity. Each dimension has four components whose underlying concepts are related. Each component is further defined by a set of three to five outcome indicators. The index is subsequently hierarchically structured, based on three dimensions, twelve components and 51 indicators. The indicators are measured on a quantitative scale.

Data availability. All data used to calculate the 2018 SPI are the most recent available as of June 2018, where most of the data are originating from 2017 or 2016. No data before 2008 has been collected. All data are publicly available deriving from diverse international organizations such as the United Nations (UN), The Institute for Health Metrics and Evaluation (IHME), Varieties of Democracy (V-Dem), Freedom House and many others (listed in the SPI Methodology report). Table 1 offers summary statistics for the SPI indicators. The tables are separated by the three dimensions. In the presented tables some preliminary imputations and data transformations made by SPI are comprised and are summarized in the following seven points:

1. Top boundary values are set to four indicators to normalize the indicator distributions from influential observations. Adult literacy rate is capped at 99%[3] (ind. 17), secondary school enrollment ratio is capped at 100%[4] (ind. 19), mobile telephone subscriptions is capped at 100 subscriptions per 100 people (ind. 22) and greenhouse gas emissions is capped at 1955,5 CO_2 equivalents per GDP-PPP (ind. 32).
2. A bottom boundary value is set to zero for the indicator measuring political rights (ind. 34). However, of the 146 ranked countries no country score was below zero.
3. Logarithmic transformations of two indicators. The indicators measuring homicide rate (ind. 16) and globally ranked universities (ind. 50) contain extreme values. In an earlier version of the SPI 2018, JRC recommended the logarithmic transformation of these indicators so their distributions become less skewed. They are transformed to avoid that the extreme values are becoming unintended benchmarks and introduce bias in the aggregation with other indicators.
4. The indicator assessing gender parity in secondary enrollment (ind. 20) is transformed as the absolute distance from one, to acknowledge the lack of parity for both boys and girls in secondary enrollment across countries.
5. The indicator evaluating access to independent media (ind. 24) is missing for four ranked countries and index developers averaged the imputations across five years to get smoother estimates.
6. The indicator measuring acceptance of gays and lesbians (ind. 42) is missing for 19 ranked countries and are imputed using regional groupings made by research from the civil rights organization, the Human Rights Campaign (HRC).

[3] The literacy rate in developed countries is often over 99% but is capped at 99% by the index developers.
[4] UNESCO's enrolment ratio includes both over- and underage children that result in results with ratios over 100%.

7. The indicator assessing adult literacy rate (ind.17) is missing for 21 developed countries and is imputed with 99% based on SPI qualitative research.

The data coverage (when the tailor made imputations in points 5-7 above are included) is generally very good to excellent. 37 indicators are complete and have no missing values while 11 indicators have less than six percent missing values. The remaining three indicators have 10% or more missing values and these are the indicators assessing quality electricity supply (ind. 11, 11% missing), adult literacy rate (ind. 17, 13% missing) and average years of tertiary schooling (ind. 48, 16% missing). For the remaining missing values, the SPI developers have used regression imputation to predict the missing values and have regressed each indicator on the other indicators within its pertaining component.

For a majority of the indicators (34 cases) the desirable direction of social progress is positive. A higher indicator value means a higher social progress, whereas for the remaining 17 indicators the opposite holds true. For example, higher values are desirable for the access to quality healthcare (ind. 29), whilst lower values are desirable for greenhouse gas emissions (ind. 32).

Table 1. Summary statistics of SPI indicators (raw data) separated by dimension.

Dimension	Component	Indicator		Direction	Nr of obs	Missing obs (%)	Mean	Stand Dev	Median	Upper boundary	Lower boundary	Range	Skewness	Kurtosis
D1: Basic Human Needs	C1: Nutrition and Basic Medical Care	Undernourishment (% of pop.)	ind.01	-	138	5%	10,6	10,7	5,6	2,5	58,6	56,1	1,8	3,4
		Maternal mortality rate (deaths/100,000 live births)	ind.02	-	146	0%	129,7	161,4	47,6	2,2	802,0	799,7	1,4	1,5
		Child mortality rate (deaths/1,000 live births)	ind.03	-	146	0%	30,6	31,0	16,2	2,1	127,3	125,2	1,2	0,4
		Child stunting (% of children)	ind.04	-	146	0%	17,0	13,5	12,8	1,0	52,4	51,4	0,7	-0,7
		Deaths from infectious diseases (deaths/100,000)	ind.05	-	146	0%	172,8	228,2	51,6	6,3	1209,9	1203,6	1,8	3,3
	C2: Water and Sanitation	Access to at least basic drinking water (% of pop.)	ind.06	+	146	0%	85,6	17,8	93,7	100,0	19,3	80,7	-1,3	1,0
		Access to piped water (% of pop.)	ind.07	+	144	1%	71,4	29,0	85,6	100,0	3,7	96,3	-0,6	-1,1
		Access to at least basic sanitation facilities (% of pop.)	ind.08	+	146	0%	72,6	29,9	86,8	100,0	7,1	92,9	-0,8	-0,9
		Rural open defecation (% of pop.)	ind.09	-	137	6%	15,4	22,4	2,3	0,0	88,7	88,7	1,5	1,4
	C3: Shelter	Access to electricity (% of pop.)	ind.10	+	146	0%	81,3	28,0	100,0	100,0	7,6	92,4	-1,3	0,1
		Quality of electricity supply (1=low; 7=high)	ind.11	+	130	11%	4,6	1,5	4,8	6,9	1,2	5,7	-0,3	-1,0
		Household air pollution attributable deaths (deaths/100,000)	ind.12	-	146	0%	53,8	66,3	18,7	0,0	262,7	262,7	1,1	0,2
	C4: Personal Safety	Political killings and torture (0=low freedom; 1=high freedom)	ind.13	+	146	0%	0,7	0,3	0,8	1,0	0,1	0,9	-1,0	-0,1
		Perceived criminality (1=low; 5=high)	ind.14	-	138	5%	3,2	1,0	3,0	1,0	5,0	4,0	-0,2	-0,4
		Traffic deaths (deaths/100,000)	ind.15	-	146	0%	18,6	11,1	16,9	3,8	59,1	55,3	1,2	1,6
		Homicide rate (deaths/100,000)	ind.16	-	146	0%	1,6	0,9	1,5	0,2	4,7	4,5	0,7	0,4

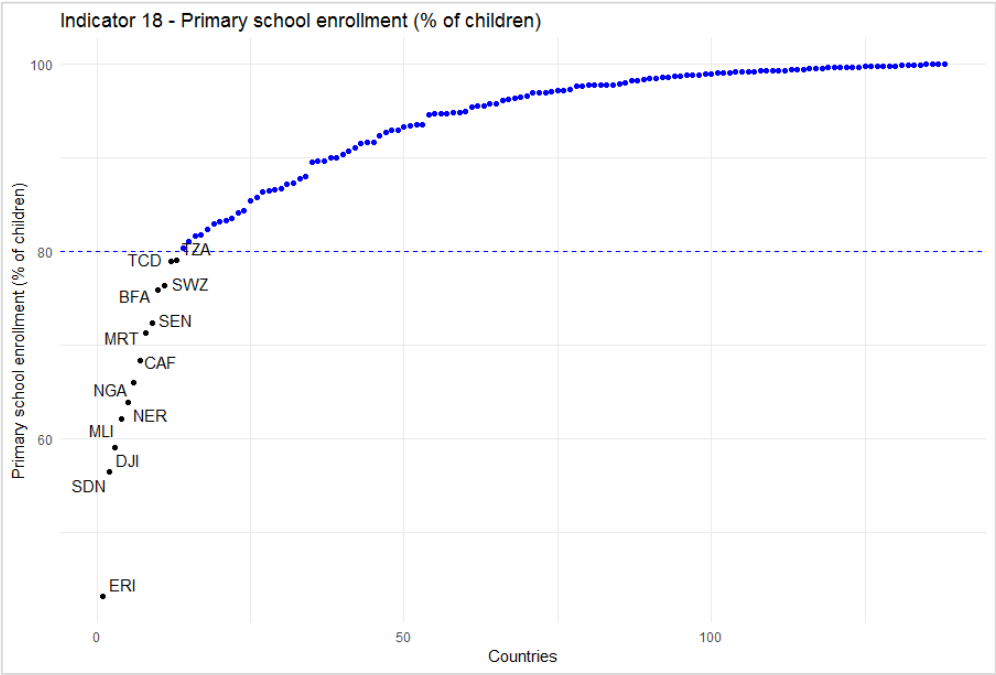
Dimension	Component	Indicator		Direction	Nr of obs	Missing obs (%)	Mean	Stand Dev	Median	Upper boundary	Lower boundary	Range	Skewness	Kurtosis
D2: Foundations of Wellbeing	C5: Access to Basic Knowledge	Adult literacy rate (% of pop. aged 15+)	ind.17	+	127	13%	83,0	21,2	94,2	99,0	15,5	83,5	-1,3	0,7
		Primary school enrollment (% of children)	ind.18	+	138	5%	92,3	10,2	96,6	100,0	43,1	56,9	***-2,1***	***5,1***
		Secondary school enrollment (% of children)	ind.19	+	146	0%	78,8	24,4	89,3	100,0	15,4	84,6	-0,8	-0,8
		Gender parity in secondary enrollment (girls/boys)	ind.20	-	145	1%	0,1	0,1	0,0	0,0	0,5	0,5	2,0	4,2
		Access to quality education (0=unequal; 4=equal)	ind.21	+	146	0%	2,2	1,1	2,4	3,9	0,2	3,7	-0,1	-1,3
	C6: Access to Info & Commun	Mobile telephone subscriptions (subscriptions/100 people)	ind.22	+	146	0%	88,8	19,5	100,0	100,0	10,2	89,8	-1,9	2,7
		Internet users (% of pop.)	ind.23	+	146	0%	49,5	28,7	52,3	98,2	1,2	97,1	0,0	-1,3
		Access to online governance (0=low; 1=high)	ind.24	+	146	0%	0,5	0,3	0,6	1,0	0,0	1,0	-0,2	-1,0
		Access to independent media (% of pop.)	ind.25	+	146	0%	65,7	23,4	63,6	100,0	14,5	85,5	-0,2	-1,0
	C7: Health and Wellness	Life expectancy at 60 (years)	ind.26	+	146	0%	20,2	3,1	20,3	26,3	12,5	13,8	-0,1	-0,9
		Non-communicable disease deaths between the ages of 30 and 70 (deaths/100,000)	ind.27	-	146	0%	411,7	186,3	387,1	168,9	1170,2	1001,3	1,6	4,0
		Access to essential health services (0=none; 100=full coverage)	ind.28	+	146	0%	61,4	13,7	62,6	86,2	30,1	56,1	0,0	-1,0
		Access to quality healthcare (0=unequal; 4=equal)	ind.29	+	146	0%	2,2	1,1	2,3	3,9	0,2	3,6	-0,1	-1,3
	C8: Environmental Quality	Outdoor air pollution attributable deaths (deaths/100,000)	ind.30	-	146	0%	60,1	38,6	51,8	6,8	207,7	200,9	0,8	0,4
		Wastewater treatment (0=no treatment; 100=fully treated)	ind.31	+	146	0%	57,5	38,3	72,3	100,0	0,0	100,0	-0,5	-1,4
		Greenhouse gas emissions (CO2 equivalents per GDP)	ind.32	-	144	1%	499,6	324,9	384,5	102,8	1955,5	1852,7	1,9	5,0
		Biome protection (% of biomes)	ind.33	+	146	0%	11,5	5,5	13,2	17,0	0,1	16,9	-0,6	-1,1

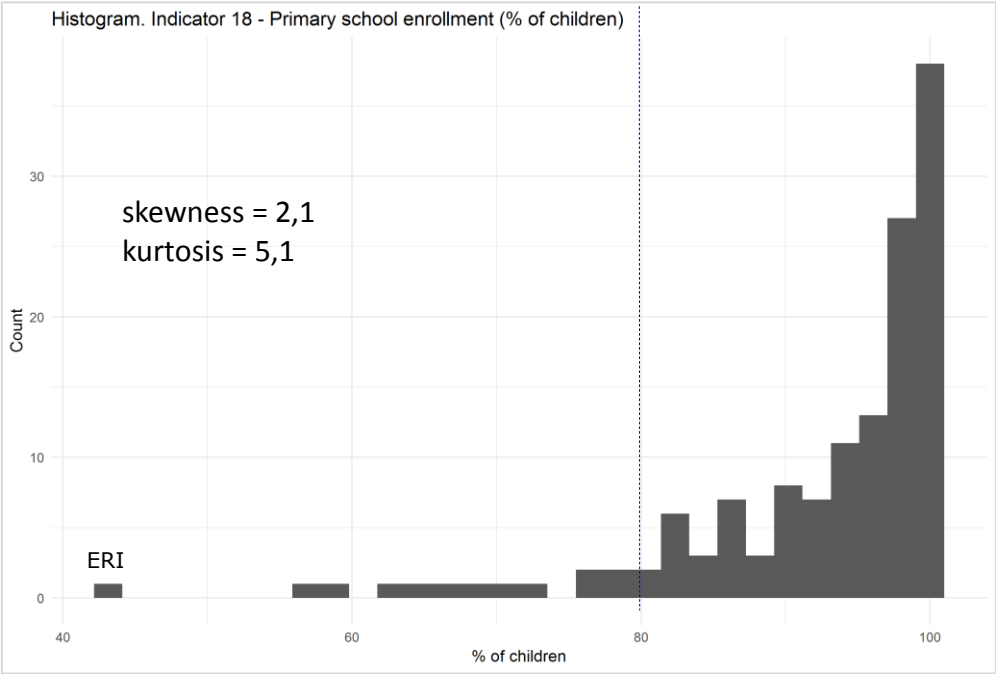
Dimension	Component	Indicator		Direction	Nr of obs	Missing obs (%)	Mean	Stand Dev	Median	Upper boundary	Lower boundary	Range	Skewness	Kurtosis
D3: Opportunity	C9: Personal Rights	Political rights (0=no rights; 40=full rights)	ind.34	+	146	0%	24,0	12,8	26,0	40,0	0,0	40,0	-0,4	-1,2
		Freedom of expression (0=no freedom; 1=full freedom)	ind.35	+	146	0%	0,7	0,3	0,8	1,0	0,0	1,0	-1,0	-0,2
		Freedom of religion (0=no freedom; 4=full freedom)	ind.36	+	146	0%	3,2	0,8	3,5	4,0	0,2	3,7	-1,6	2,1
		Access to justice (0=non-existent; 1=observed)	ind.37	+	146	0%	0,7	0,2	0,7	1,0	0,1	0,9	-0,5	-0,9
		Property rights for women (0=no right; 5=full rights)	ind.38	+	146	0%	4,0	1,0	4,3	4,9	0,8	4,1	-1,3	0,8
	C10: Personal Freedom and Choice	Vulnerable employment (% of employees)	ind.39	-	146	0%	36,1	26,2	31,8	0,2	94,5	94,3	0,6	-0,9
		Early Marriage (% of women)	ind.40	-	137	6%	0,1	0,1	0,1	0,0	0,6	0,6	1,3	1,4
		Satisfied demand for contraception (% of women)	ind.41	+	143	2%	65,1	19,1	71,1	94,8	19,7	75,1	-0,6	-0,8
		Corruption (0=high; 100=low)	ind.42	+	145	1%	44,4	19,1	39,0	89,0	15,0	74,0	0,8	-0,5
	C11: Inclusiveness	Acceptance of gays and lesbians (0=low; 100=high)	ind.43	+	146	0%	0,3	0,3	0,2	0,9	0,0	0,9	0,9	-0,4
		Discrimination and violence against minorities (1=low; 10=high)	ind.44	-	146	0%	6,1	2,1	6,3	1,1	10,0	8,9	-0,2	-0,8
		Equality of political power by gender (0=unequal power; 4=equal power)	ind.45	+	146	0%	2,0	0,7	2,0	3,6	0,3	3,3	-0,2	-0,3
		Equality of political power by socioeconomic position (0=unequal power; 4=equal power)	ind.46	+	146	0%	2,0	0,8	2,1	3,7	0,1	3,5	-0,3	-0,5
		Equality of political power by social group (0=unequal power; 4=equal power)	ind.47	+	146	0%	2,3	0,9	2,4	3,8	0,1	3,7	-0,5	-0,5
	C12: Access to Advanced Education	Years of tertiary schooling	ind.48	+	123	16%	0,6	0,5	0,5	1,9	0,0	1,9	0,9	0,3
		Women's average years in school	ind.49	+	146	0%	9,9	4,2	10,8	15,7	1,0	14,7	-0,5	-1,1
		Number of globally ranked universities (points)	ind.50	+	146	0%	1,3	1,5	0,7	5,9	0,0	5,9	0,9	-0,3
		Percent of tertiary students enrolled in globally ranked universities	ind.51	+	145	1%	0,2	0,2	0,0	0,8	0,0	0,8	1,2	0,6

Notes: Raw data refer to year 2018. Practical JRC rule for outlier detection: Indicators with |skewness|>2 and kurtosis >3. N=146 countries.

Source: European Commission, Joint Research Centre, 2018.

Figure 1. Eritrea's performance in primary school enrollment (% of children).





Source: European Commission, Joint Research Centre, 2018.

Outlier detection. Potentially problematic indicators that could bias the overall index results were identified on the basis of two measures related to the shape of the distributions: skewness and kurtosis. A practical rule suggested by the JRC is that country values should be considered and possibly treated if the indicators have absolute

skewness greater than 2,0 and kurtosis greater than 3,5[5]. In an earlier version of the 2018 SPI, JRC recommended data transformations to the following three indicators (whose distributions were skewed) measuring homicide rate (ind. 16), primary school enrollment rate (ind. 18) and globally ranked universities (ind. 50). For the final version, two of these indicators have been (effectively) transformed (ind. 16 and ind. 50 have been log transformed).

As shown in Table 1 and Figure 1, the indicator measuring primary school enrollment rate (ind. 18) is still possibly problematic where the low value of 43% for Eritrea is deviating. The skewness is high and negative (-2,1), which implies that the distribution of the data is skewed to the left (*i.e.* negatively skewed, as seem in the histogram in Figure 1). The kurtosis is high (5,1) and considerably higher than for a normal distribution (3). It has a distribution with larger tails (*i.e. leptokurtic distribution*). However, the deviating value does not seem to be erroneous since it follows smoothly the time series for Eritrea of the years 2014-2018 (figure not shown) and the indicator is left untransformed. Moreover, there are only countries in Africa (13 countries), with primary school enrollment rates below 80%. Preferably the indicator should be monitored in the coming releases of the index and eventually a data winsorization[6] may be applied.

Normalization. Best- and worst-case scenarios are set to provide actual boundaries on both ends of the scale that are based on theoretical or historical values. For the index calculation the data are normalized twice. First, the raw data for the 51 SPI indicators are put on a common scale using z-score standardization[7]. Data-driven weights based on principal components analysis (PCA) are in addition retrieved. The values of the SPI components are calculated as the aggregation of the indicators within each component multiplied with each scaled principal component weight. Secondly, the components values are normalized using the min-max normalization method[8] on a scale of 0 and 100 to retrieve the component scores. The use of PCA-weights is further explained in Section 4.

Aggregation. Arithmetic averaging is used in all levels to build the SPI; at the first aggregation level (from indicators to components), the second (from components to dimensions) and at the third and last aggregation level (from dimensions to an overall index). This means that each dimension is the arithmetic average of the four components that make up that dimension and the overall index is calculated as the arithmetic average over the three dimensions. Arithmetic averages are easy interpret and allow perfect compensability between the variables, whereby a high score on one variable can fully offset low scores in other variables.

Weights. The SPI developers opted to use objective weights at the lowest aggregation level and chose the intrinsic weights from the PCA factor loadings. In the following two aggregation levels equal weights have been assigned to the components and the dimensions. Figure 2 illustrates the different weights, aggregation methods and normalization methods employed in the SPI framework.

---

[5] Groeneveld and Meeden (1984) set the criteria for absolute skewness above 1 and kurtosis above 3.5. The skewness criterion was relaxed in the SPI case after having conducted ad-hoc tests in the SPI 2014-2018 timeseries.

[6] By winsorization, one converts the value(s) of data points that are considered to be too high (low) to a value of the highest (lowest) data point not considered to be an outlier.

[7] With the z-scores method, a country's score for each indicator is calculated by subtracting the average value (across all countries) and dividing by the standard deviation.

[8] With the min-max normalization method, a country's score for each indicator is calculated by subtracting the minimum value (or lower bound) across all countries and dividing by the difference between the maximum and minimum values (or upper and lower bounds).

Figure 2. Normalization, aggregation methods and weights used in the SPI framework.

1 Indicator level				2 Component level				3 Dimension level			4 Index level
Indicator	Normal-isation 1	PCA weights	Artihm. Mean	Component	Normal-isation 2	Equal Weights	Artihm. Mean	Dimension	Equal Weights	Artihm. Mean	Index
ind 1	z-score standardisation	0,18		C1. Nutrition and Basic Medical Care	Min-max normalisation	0,25		D1. Basic Human Needs	0,33		Social Progress Index
ind 2		0,21									
ind 3		0,21									
ind 4		0,20									
ind 5		0,20									
ind 6		0,26		C2. Water and Sanitation		0,25					
ind 7		0,24									
ind 8		0,27									
ind 9		0,23									
ind 10		0,33		C3. Shelter		0,25					
ind 11		0,32									
ind 12		0,34									
ind 13		0,25		C4. Personal Safety		0,25					
ind 14		0,27									
ind 15		0,24									
ind 16		0,24									
ind 17		0,22		C5. Access to Basic Knowledge		0,25		D2. Foundations of Wellbeing	0,33		
ind 18		0,20									
ind 19		0,22									
ind 20		0,18									
ind 21		0,18									
ind 22		0,23		C6. Access to Information and Communications		0,25					
ind 23		0,27									
ind 24		0,25									
ind 25		0,24									
ind 26		0,27		C7. Health and Wellness		0,25					
ind 27		0,24									
ind 28		0,27									
ind 29		0,23									
ind 30		0,29		C8. Environmental Quality		0,25					
ind 31		0,27									
ind 32		0,26									
ind 33		0,17									
ind 34		0,21		C9. Personal Rights		0,25		D3. Opportunity	0,33		
ind 35		0,21									
ind 36		0,19									
ind 37		0,20									
ind 38		0,19									
ind 39		0,26		C10. Personal Freedom and Choice		0,25					
ind 40		0,26									
ind 41		0,23									
ind 42		0,25									
ind 43		0,19		C11. Inclusiveness		0,25					
ind 44		0,18									
ind 45		0,21									
ind 46		0,21									
ind 47		0,21									
ind 48		0,26		C12. Access to Advanced Education		0,25					
ind 49		0,26									
ind 50		0,25									
ind 51		0,23									

Source: European Commission, Joint Research Centre, 2018.

3 Statistical coherence of the SPI

The reliability of the Social Progress Index depends - among other things – on the degree of coherence between the conceptual framework and the statistical structure of the data. The more the SPI conceptual framework encompasses the statistical structure, the higher the reliability of the SPI will be. The coherence of the SPI framework was assessed analyzing the extent to which the SPI indicators can explain a sufficient amount of variation in the aggregated scores (be those components, dimensions or the overall index) by means of correlation, cross-correlation and principal component analysis (PCA). Given that the present statistical analysis of the Social Progress Index is in part, though not exclusively, based on correlations, the correspondence of the SPI to a real-world phenomenon needs to be critically addressed by experts in the field because "correlations need not necessarily represent the real influence of the individual indicators on the phenomenon being measured"[9].

Succinctly, the argument is that the validity of the SPI framework relies on the combination of both statistical and conceptual soundness. In this respect, the SPI framework has been developed following an iterative process that went back and forth between the theoretical understandings of an outcome-based measure of countries wellbeing on the one hand, and data observations on the other. Starting with the simplest approach, correlation and cross-correlation analysis was used to assess to what extent the data collected support the SPI conceptual framework.

The SPI framework is characterized by strong to high correlations. The highest correlations are found in the dimension Foundations of Wellbeing; with a correlation of 0,92 between the indicators measuring education equality (ind. 21) and health equality (ind. 29)[10] and followed by a correlation of 0,91 between the indicators assessing internet users in the population (ind. 23) and access to essential health services (ind. 28). The indicators within any component exhibit generally strong correlations between them (majority over 0,60), apart from the indicator on the percentage of biomes (naturally occurring community of flora and fauna) in protected areas (ind. 33) in the environmental quality component (C8), where the correlations are less than 0,30 or insignificant.

A more detailed analysis of the SPI correlation structure within and between the 12 components shows that all indicators are positively and strongly correlated with their pertaining components. About 75% (37/51) of the correlations are greater than 0,80 and of those 25% (12/51) are very strong with correlations over 0,92. These strong correlations are found in all three dimensions but most cases are found in the first two.

[9] OECD & EC JRC (2008).
[10] The source for both indicators is Varieties of Democracy (V-Dem) project, https://www.v-dem.net/en/

Table 2. Statistical coherence in the SPI framework. Correlations between indicators and other SPI aggregates.

Dimension 1: Basic Human Needs

Dimension	Component	Indicator		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	D1	D2	D3	Index
D1: Basic Human Needs	C1: Nutrition and Basic Medical Care	Undernourishment	ind.01	0,85	0,75	0,77	0,54	0,67	0,73	0,69	0,59	0,43	0,64	0,44	0,67	0,80	0,72	0,64	0,76
		Maternal mortality rate	ind.02	**0,94**	0,88	0,88	0,54	0,88	0,80	0,76	0,77	0,40	0,80	0,39	0,72	0,90	0,86	0,68	0,85
		Child mortality rate	ind.03	**0,93**	0,87	0,88	0,58	0,90	0,82	0,77	0,78	0,42	0,83	0,43	0,77	0,91	0,88	0,72	0,87
		Child stunting	ind.04	0,88	0,83	0,84	0,65	0,82	0,82	0,82	0,76	0,52	0,81	0,58	0,80	0,88	0,87	0,79	0,88
		Deaths from infectious diseases	ind.05	0,90	0,79	0,82	0,51	0,75	0,72	0,69	0,64	0,32	0,66	0,28	0,63	0,84	0,75	0,56	0,75
	C2: Water and Sanitation	Access to basic drinking water	ind.06	0,88	**0,93**	0,89	0,55	0,83	0,83	0,75	0,73	0,46	0,76	0,45	0,72	0,91	0,85	0,70	0,86
		Access to piped water	ind.07	0,85	0,89	0,87	0,54	0,77	0,76	0,79	0,75	0,43	0,78	0,49	0,71	0,88	0,82	0,71	0,84
		Access to basic sanitation facilities	ind.08	0,91	**0,96**	0,92	0,51	0,86	0,82	0,78	0,73	0,33	0,83	0,36	0,78	**0,93**	0,86	0,68	0,86
		Rural open defecation	ind.09	0,68	0,85	0,69	0,37	0,72	0,67	0,58	0,59	0,25	0,64	0,23	0,62	0,74	0,69	0,51	0,68
	C3: Shelter	Access to electricity	ind.10	0,90	0,89	**0,93**	0,46	0,82	0,79	0,71	0,70	0,33	0,74	0,34	0,70	0,89	0,81	0,62	0,81
		Quality of electricity supply	ind.11	0,80	0,81	0,91	0,66	0,81	0,81	0,85	0,78	0,41	0,83	0,54	0,80	0,89	0,87	0,76	0,88
		Household air pollution deaths	ind.12	0,89	0,88	**0,95**	0,49	0,82	0,78	0,83	0,78	0,35	0,80	0,38	0,74	0,90	0,86	0,67	0,85
	C4: Personal Safety	Political killings and torture	ind.13	0,54	0,48	0,48	0,78	0,51	0,64	0,58	0,57	0,85	0,54	0,77	0,46	0,61	0,62	0,77	0,69
		Perceived criminality	ind.14	0,46	0,43	0,46	0,82	0,51	0,49	0,52	0,41	0,43	0,56	0,57	0,47	0,58	0,52	0,58	0,58
		Traffic deaths	ind.15	0,53	0,43	0,47	0,73	0,50	0,51	0,63	0,58	0,61	0,44	0,54	0,52	0,57	0,59	0,62	0,62
		Homicide rate	ind.16	0,40	0,32	0,37	0,74	0,45	0,44	0,47	0,37	0,25	0,41	0,34	0,49	0,48	0,46	0,44	0,48

Dimension 2: Foundations of Wellbeing

Dimension	Component	Indicator		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	D1	D2	D3	Index
D2: Foundations of Wellbeing	C5: Access to Basic Knowledge	Adult literacy rate	ind.17	0,84	0,82	0,83	0,46	0,91	0,76	0,70	0,75	0,36	0,81	0,39	0,75	0,83	0,84	0,68	0,82
		Primary school enrollment	ind.18	0,65	0,67	0,61	0,46	0,80	0,66	0,54	0,60	0,38	0,60	0,37	0,59	0,66	0,69	0,57	0,67
		Secondary school enrollment	ind.19	0,91	0,90	0,90	0,56	**0,93**	0,83	0,77	0,75	0,45	0,83	0,49	0,79	0,91	0,88	0,75	0,89
		Gender parity in secondary enrollment	ind.20	0,59	0,55	0,57	0,46	0,70	0,51	0,48	0,52	0,30	0,51	0,25	0,48	0,60	0,59	0,45	0,57
		Access to quality education	ind.21	0,62	0,59	0,63	0,69	0,73	0,66	0,73	0,58	0,50	0,69	0,63	0,64	0,69	0,72	0,72	0,74
	C6: Access to Info & Commun	Mobile telephone subscriptions	ind.22	0,69	0,70	0,67	0,44	0,67	0,78	0,57	0,59	0,48	0,59	0,33	0,54	0,70	0,71	0,57	0,69
		Internet users	ind.23	0,87	0,87	0,88	0,66	0,84	**0,92**	0,86	0,80	0,52	0,88	0,56	0,88	0,91	**0,92**	0,83	**0,93**
		Access to online governance	ind.24	0,76	0,72	0,76	0,56	0,75	0,88	0,74	0,70	0,49	0,75	0,48	0,81	0,78	0,83	0,75	0,82
		Access to independent media	ind.25	0,64	0,62	0,61	0,63	0,63	0,85	0,70	0,68	0,74	0,68	0,68	0,66	0,68	0,78	0,81	0,79
	C7: Health and Wellness	Life expectancy at 60 years	ind.26	0,81	0,78	0,83	0,63	0,73	0,81	**0,96**	0,82	0,50	0,75	0,55	0,77	0,84	0,89	0,75	0,87
		Non-communicable disease deaths	ind.27	0,54	0,53	0,62	0,47	0,49	0,58	0,85	0,66	0,36	0,50	0,43	0,53	0,60	0,69	0,53	0,63
		Access to essential health services	ind.28	0,89	0,85	0,88	0,73	0,86	0,89	**0,96**	0,83	0,51	0,87	0,59	0,87	**0,92**	**0,95**	0,83	**0,94**
		Access to quality healthcare	ind.29	0,72	0,69	0,73	0,73	0,75	0,74	0,83	0,64	0,52	0,76	0,66	0,71	0,79	0,79	0,77	0,82
	C8: Environmental Quality	Outdoor air pollution deaths	ind.30	0,73	0,70	0,73	0,59	0,74	0,71	0,79	0,88	0,58	0,74	0,60	0,70	0,76	0,83	0,77	0,82
		Wastewater treatment	ind.31	0,78	0,79	0,81	0,52	0,77	0,82	0,75	0,83	0,44	0,77	0,47	0,81	0,81	0,85	0,73	0,83
		Greenhouse gas emissions	ind.32	0,57	0,54	0,59	0,38	0,54	0,55	0,56	0,77	0,32	0,58	0,32	0,47	0,58	0,64	0,50	0,60
		Biome protection	ind.33				0,32		0,28	0,30	0,48	0,40		0,40	0,24		0,34	0,36	0,31

Dimension 3: Opportunity

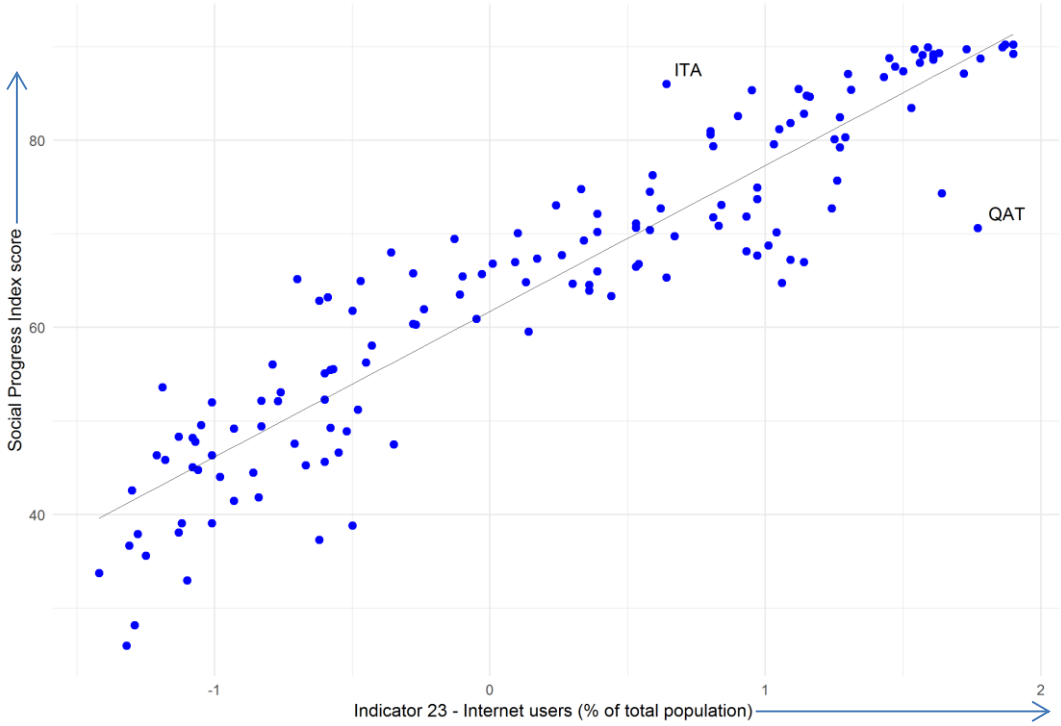
Dimension	Component	Indicator		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	D1	D2	D3	Index
D3: Opport-unity	C9: Personal Rights	Political rights	ind.34	0,46	0,41	0,41	0,65	0,47	0,64	0,52	0,59	**0,95**	0,53	0,79	0,51	0,52	0,60	0,81	0,67
		Freedom of expression	ind.35	0,30	0,25	0,23	0,55	0,30	0,51	0,38	0,46	**0,92**	0,36	0,73	0,35	0,35	0,45	0,69	0,51
		Freedom of religion	ind.36				0,37	0,25	0,35	0,23	0,37	0,78	0,22	0,63	0,23		0,33	0,55	0,36
		Access to justice	ind.37	0,53	0,47	0,47	0,79	0,54	0,65	0,60	0,55	0,85	0,60	0,73	0,54	0,60	0,63	0,80	0,70
		Property rights for women	ind.38	0,57	0,48	0,49	0,61	0,54	0,65	0,54	0,54	0,83	0,47	0,60	0,50	0,58	0,61	0,71	0,66
	C10: Personal Freedom and Choice	Vulnerable employment	ind.39	0,79	0,77	0,81	0,46	0,73	0,73	0,71	0,66	0,31	0,87	0,42	0,76	0,79	0,76	0,69	0,78
		Early Marriage	ind.40	0,73	0,73	0,73	0,50	0,79	0,74	0,62	0,67	0,38	0,86	0,43	0,71	0,75	0,76	0,69	0,77
		Satisfied demand for contraception	ind.41	0,60	0,62	0,64	0,38	0,64	0,58	0,57	0,63	0,36	0,75	0,47	0,61	0,63	0,65	0,64	0,66
		Corruption	ind.42	0,63	0,62	0,66	0,78	0,65	0,77	0,77	0,70	0,64	0,83	0,76	0,74	0,73	0,78	0,86	0,82
	C11: Inclusive-ness	Acceptance of gays and lesbians	ind.43	0,51	0,49	0,53	0,52	0,50	0,61	0,63	0,62	0,57	0,64	0,77	0,63	0,56	0,64	0,76	0,67
		Discrimination/violence against minorities	ind.44	0,34	0,29	0,31	0,57	0,39	0,40	0,44	0,40	0,53	0,46	0,72	0,36	0,40	0,43	0,59	0,49
		Equality of political power by gender	ind.45	0,31	0,28	0,33	0,58	0,41	0,48	0,49	0,45	0,69	0,47	0,83	0,44	0,40	0,49	0,70	0,55
		Equality of political power by socioecon. pos.	ind.46	0,39	0,35	0,39	0,65	0,46	0,49	0,53	0,51	0,65	0,53	0,84	0,49	0,47	0,54	0,73	0,60
		Equality of political power by social group	ind.47	0,30	0,26	0,23	0,55	0,30	0,42	0,36	0,40	0,76	0,35	0,79	0,36	0,35	0,40	0,66	0,48
	C12: Access to Advanced Education	Years of tertiary schooling	ind.48	0,69	0,64	0,67	0,58	0,70	0,72	0,69	0,62	0,43	0,69	0,49	0,86	0,71	0,73	0,73	0,75
		Women's average years in school	ind.49	0,83	0,83	0,82	0,58	0,90	0,83	0,74	0,77	0,49	0,84	0,51	0,87	0,85	0,87	0,80	0,87
		Number of globally ranked universities	ind.50	0,59	0,58	0,62	0,43	0,57	0,66	0,63	0,59	0,31	0,66	0,41	0,84	0,62	0,66	0,65	0,67
		% tertiary stud. globally ranked universities	ind.51	0,58	0,58	0,61	0,54	0,57	0,68	0,68	0,65	0,44	0,69	0,54	0,80	0,64	0,69	0,73	0,71

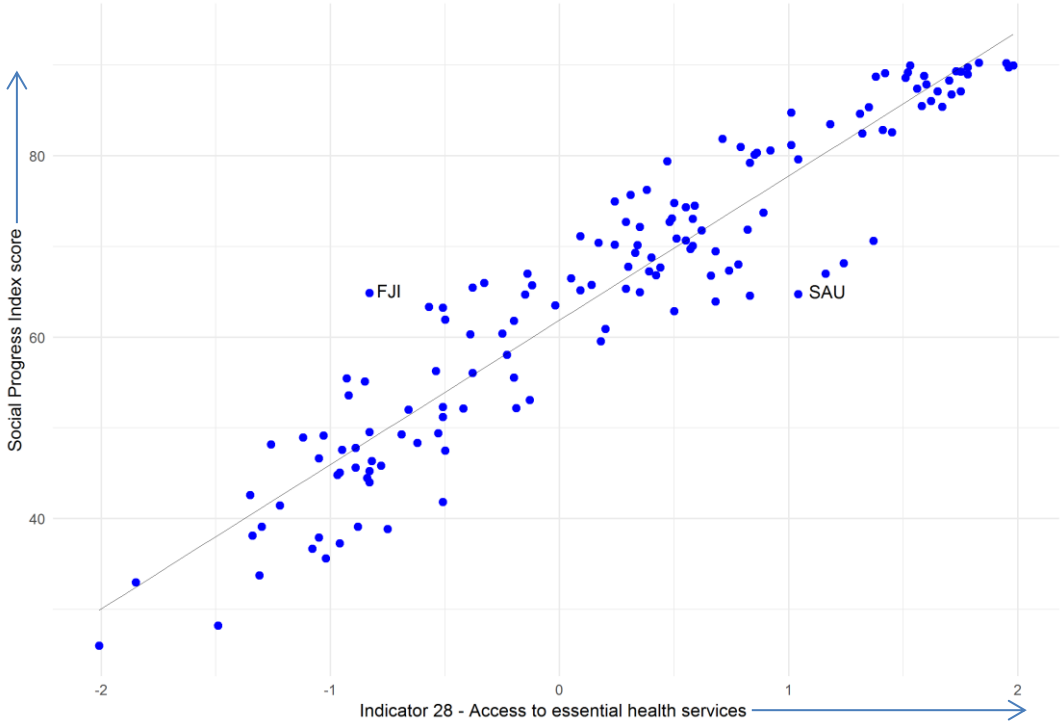
Notes: Numbers represent the Pearson correlations coefficients between the SPI aggregates (components, dimensions, index) and the underlying indicators (for 146 countries). Correlations that are not significant at the significance level of α = 0,01 are left blank (critical value of 0.21). Grey boxes show the conceptual grouping of the indicators. Very strong correlations (i.e. Pearson correlation coefficients greater than 0,92) are marked in bold.

Source: European Commission, Joint Research Centre, 2018.

All the indicators are correlating positively and significantly with the overall index. Furthermore, in a majority of cases the correlations are well over 0,7, which suggests that at least 50% of the total variation is captured by the overall index, SPI. The universal health coverage index indicator (ind. 28) and the indicator estimating the number of internet users (ind. 23) in the second dimension Foundations of Wellbeing are correlating the highest with the SPI (correlations 0,94 and 0,93, respectively). The scatterplots in Figure 3 show the very high linear relationships between the SPI scores and the standardized values (z-scores) of indicators 23 and 28. Albeit the very high correlations which indicate that the indicators and SPI are measuring similar concepts, they prove to contain useful information as well. The share of internet users (ind.23, top panel, Figure 3) in Italy (ITA) is lower than what would be expected for an overall SPI score of 86 points. In Qatar (QAT), on the other hand the opposite holds true, there are more internet users in the population than what would be expected with a SPI score of 76 points. In the case of access to essential health services (ind.28, bottom panel, Figure 3), Fiji (FJI) and Saudi Arabia (SAU) have similar scores of SPI (65 points) but very different scores of health service availability (with higher scores for the latter). The lowest correlation with the SPI, is found for the indicator measuring the percentage biomes in protected areas (ind.31), which is also found in the second dimension (the correlation is low, 0.31, but still significant).

Figure 3. Scatter plots of SPI 2018 versus indicator 23 and 28 scores





Notes: The dots represent the country scores for 146 countries for Social Progress Index versus z-scores of the indicator estimating the number of internet users (ind.23, top panel) and the essential health services (ind.18, bottom panel). Black line denotes the regression line.

Source: European Commission, Joint Research Centre, 2018.

At the next aggregation level, there are three out of the 12 components that are better associated with one or both other dimensions than to where they have been assigned (Table 3). The Personal Safety (C4) component is better correlated with the Opportunity dimension than to its assigned Basic Human Needs dimension. This component is also correlating better with all other (excluding C8) components compared to its own dimension's components. The Personal Freedom and Choice component (C10) is marginally better associated with the first two dimensions compared to the Opportunity dimension. The Access to Advanced Education (C12) is correlating slightly higher with the Foundations of Wellbeing dimension compared to the Opportunity dimension.

Generally, this is an undesired situation that the allocation of the components to a specific dimension is not consistent from the conceptual to the statistical perspective. In the SPI framework, on the other hand, some components are very interrelated because of strong connections between the dimensions. Take for example the two related knowledge components, Access to Basic Knowledge (C5) in the Basic Human Needs dimension and the component Access to Advanced Education (C12) in the Opportunity dimension. The indicator measuring women's average years in education (ind. 49) is marginally more correlated with the Basic Knowledge (C5) compared to its Advanced Education (C12) component. This is not surprising since the indicator includes the number of years in school from basic to advanced education[11].

At the highest aggregation level, the three dimensions Basic Human Needs, Foundations of Wellbeing and Opportunity correlate very strongly and in a balanced way with the SPI, with pairwise correlations between 0,93 and 0,98. The Foundations of Wellbeing is correlating the strongest with the index. As anticipated the three dimensions are also correlating strongly with each other, with pairwise correlations between 0,81 and 0,95. This result may flag potential redundancy of information because the three dimensions may be measuring very similar concepts. This outcome will be further elaborated in Section 5.

Subsequently, PCA was used to confirm whether there is a single statistical dimension in each SPI aggregate (be that component, dimension or the overall index), which would give the "statistical justification" for aggregating indicators into one number. Technically, the expectation here is that there is only one principal component with an eigenvalue greater than one. Indeed, PCA results corroborate the presence of a single latent dimension in each of the 12 SPI components that captures between 57% (environmental quality component, C8) and 86% (shelter component, C3) of the total variance in the underlying indicators. Further at dimension level, PCA analysis confirms "unidimensionality" in each of the three dimensions, the single latent dimension captures 82% in the dimension Basic Human Needs, 86% in dimension Foundations of Wellbeing and 73% in dimension Opportunity of the total variance of the underlying components. Finally, the three SPI dimensions share a single statistical dimension that summarizes 92% of the total variance. This latter result supports the aggregation of three dimensions into one number.

[11] The average number of years of school attended by women between 25 and 34 years old, including primary, secondary and tertiary education.

Table 3. Statistical coherence in the SPI framework. Correlations between components and other SPI aggregates.

SPI components, dimensions and index		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	D1	D2	D3	Index
Nutrition and Basic Medical Care	C1	1,00												**0,96**	0,91	0,75	0,91
Water and Sanitation	C2	0,91	1,00											**0,95**	0,89	0,72	0,89
Shelter	C3	**0,93**	**0,93**	1,00										**0,96**	0,91	0,73	0,91
Personal Safety	C4	0,63	0,54	0,58	1,00									0,73	0,71	0,78	0,77
Access to Basic Knowledge	C5	0,89	0,87	0,88	0,64	1,00								0,91	**0,92**	0,78	0,91
Access to Information and Communications	C6	0,86	0,85	0,85	0,68	0,85	1,00							0,90	**0,95**	0,87	**0,94**
Health and Wellness	C7	0,83	0,80	0,86	0,71	0,79	0,84	1,00						0,88	**0,93**	0,81	0,91
Environmental Quality	C8	0,79	0,77	0,81	0,62	0,79	0,81	0,83	1,00					0,83	0,92	0,81	0,89
Personal Rights	C9	0,46	0,41	0,39	0,69	0,48	0,65	0,52	0,58	1,00				0,52	0,61	0,83	0,67
Personal Freedom and Choice	C10	0,83	0,83	0,85	0,64	0,85	0,85	0,81	0,80	0,51	1,00			0,88	0,89	0,87	0,91
Inclusiveness	C11	0,47	0,42	0,45	0,73	0,52	0,61	0,62	0,60	0,81	0,62	1,00		0,55	0,63	0,87	0,71
Access to Advanced Education	C12	0,79	0,78	0,80	0,63	0,81	0,85	0,81	0,78	0,49	0,85	0,58	1,00	0,83	0,87	0,86	0,89
Basic Human Needs	D1													1,00			**0,96**
Foundations of Wellbeing	D2													**0,95**	1,00		**0,98**
Opportunity	D3													0,81	0,88	1,00	**0,93**

Notes: Numbers represent the Pearson correlations coefficients between the SPI aggregates (components, dimensions, index) and the underlying indicators (for 146 countries). Correlations that are not significant at the significance level of α = 0,01 are left blank (critical value of 0.21). Grey boxes show the conceptual grouping of the indicators. Very strong correlations (i.e. Pearson correlation coefficients greater than 0,92) are marked in bold.

Source: European Commission, Joint Research Centre, 2018.

Complementary to the PCA on the current SPI structure, PCA was also performed on the whole set of 51 indicators. Not unexpectedly the number of principal components was lower than the current number of components in the SPI framework, due to the very strong correlations between the SPI aggregates. Seven latent dimensions (principal components) are retrieved which is considerably less than the current 12 components. The seven latent dimensions capture 78% of the total variance in the underlying indicators. This means that the SPI framework without losing important information may reduce the number of components, either by combining already existing ones and/or dropping some components.

Less number of components is also in line with the results from the "beyond GDP" approach by the Commission on the Measurement of Economic Performance and Social Progress[12]. The Commission identified eight[13] key dimensions that should be taken into account when defining well-being. The SPI conceptual framework has been influenced by this work. So apart from the statistical reasoning, the result from the "beyond GDP" approach, may also give conceptual justification for reducing the number of components in the SPI framework.

[12] Stiglitz, J., Sen,A., Fitoussi,J.-P., 2009. Report by the Commission on the Measurement of Economic Performance and Social Progress. In: Commission on the Measurement of Economic Performance and Social Progress , Paris, France.

[13] Material living standards (income, consumption and wealth); Health; Education; Personal activities including work; Political voice and governance; Social connections and relationships; Environment (present and future conditions); Insecurity, of an economic as well as a physical nature.

3.1 JRC recommendations based on the statistical coherence analysis

The performed statistical coherence analysis confirms the three-level structure in the SPI framework and the "unidimensionality" of all SPI aggregates (components, dimensions and index). Furthermore, all 51 indicators were found to be influential from the first aggregation level (component) up to the overall index. This is a desirable outcome as it suggests that the information content in all underlying indicators is maintained at all levels of aggregation in the SPI framework.

However, there are redundancies in the SPI framework due to high correlations between the SPI aggregates. Three components fit better statistically in other dimensions compared to where they have been allocated. The three "cross-cutting" components are the Personal Safety (C4), the Personal Freedom and Choice (C10) and Access to Advanced Education (C12) components. PCA on the whole set of 51 indicators has also shown that the number of components in the SPI framework may be reduced without losing information. Seven to nine components would suffice to explain social progress adequately. The recommendation is therefore to consider simplifying the SPI framework by:

- Reducing the number of highly correlated indicators (within the same component), either by giving them less weight or by dropping one of them. Some indicators may be mirroring the same information, therefore using only one of them could be sufficient.
- Reducing the number of components between seven and nine for instance, either by combining already existing ones and/or dropping some components.

4 Impact of modelling assumptions in the SPI

A fundamental step of the statistical analysis of a composite indicator is to assess the effect of different modelling assumptions among reasonable alternatives. In the specific case of the SPI, there are many layers which can be taken into account:

- Much research is behind the structure of the SPI, whose conceptual framework is based on experts opinion and literature review;
- the treatment of missing values is performed with statistical linear regression, within every component the missing values are estimated starting from the other variables of the component;
- the elementary indicators are sometimes capped or transformed (by logarithm) to avoid the presence of misleading outliers;
- the first normalization, at the indicators' level, is performed with the z-scores approach, in order to have variables with the same variance;
- the selection of the weights of the elementary indicators is based on the data-driven method of principal component analysis (discussed in the next section);
- the second normalization, at the component level, is performed with the min-max approach;
- the aggregation is based on weighted arithmetic mean, which is fully compensatory;
- and finally, the use of equal weights for the aggregation of components and dimensions underlines the assumption of equal importance among the constituting factors of the index.

Despite the efforts given in the building process, there is an unavoidable subjectivity (or uncertainty) in the resulting choices. This subjectivity is accounted for in the robustness assessment carried out by the JRC. The uncertainty analysis is conducted herein to allow for the analysis of the impact of some of the modelling choices on the SPI results.

The literature on this topic[14] suggests to assess the robustness basing on a Monte Carlo simulation and multi-modelling approach, assuming 'error free' data because eventual errors have been corrected in the preliminary stage of the construction. The modelling issues considered in the assessment of the SPI were the aggregation formula at the dimension level and the weights of dimensions. Regarding the method for the imputation of missing data, the developers opted for an approach based on linear regression within the components, which is well fitting according to the correlation structure of data. Some of the missing values are also imputed according to expert opinion, for these reasons the imputation method has not been included in the modelling assumptions to perturb. The normalization method proposed by the developers is strictly related to the definition of the weights in the first level of aggregation; at this level, the developers want to obtain 12 components that are representing 12 unidimensional concepts. Indeed, with the PCA approach to the definition of weights, the resulting weights are the ones that maximise the representation of the constituting variables in the components. This is a useful property of the PCA weights, but it does not assure equal representability of the constituting indicators or recognition of equal importance. This modelling choice is not included in the uncertainty analysis because a change at this level may bring uncontrollable changes to the structure of the index. Instead, a comparison between the

[14] Saisana et al., 2005; Saisana et al., 2011

PCA-weights and the equal weights has been performed; it shows that most of the weights derived from the data-driven approach are very similar to the hypothetical equal weights except for the indicators belonging to the Environmental Quality component (C8). As a consequence, it is possible to consider simplifying the SPI avoiding the z-scores normalization and the PCA step without expecting large changes in the results. This change is not mandatory and does not imply a methodological improvement. So it is up to the authors and their interpretation of the components.

Aggregation formula. Regarding the aggregation formula at the index level, as shown in Figure 2, the SPI team opted for the arithmetic averaging of the three dimensions which implies a complete compensability that allows outstanding performance in some aspects to balance for weaknesses in others and vice-versa. This approach puts at the same level countries showing high and low results with balanced countries showing average results. To assess the impact of this choice, the JRC included in the analysis both the arithmetic and the geometric mean. The comparison of the two aggregation approaches should be able to highlight the countries with unbalanced profiles because the geometric mean tends to penalize the existence of a low value, even when the other values are not so low.

Weights. Monte Carlo simulation comprised 1,000 runs of different set of weights for the three dimensions constituting the SPI: Basic Human Needs, Foundations of Wellbeing and Opportunity. The weights are the results of a random extraction based on uniform continuous distributions centred in the reference value, that is 1/3, plus or minus 25% of this value.

Two models were tested comparing the different aggregation formula obtaining a total of 2,000 runs of simulation for the SPI.

Table 4. Modelling assumptions considered in the analysis.

	Reference	**Alternative**
I. Uncertainty in the aggregation formula at the dimension level	Arithmetic average	Geometric average
II. Uncertainty intervals for the SPI dimension weights	Reference value for the weight	Distribution for robustness analysis
Basic Human Needs Dimension	0,33	U[0,27;0,40]
Foundations of Wellbeing Dimension	0,33	U[0,27;0,40]
Opportunity Dimension	0,33	U[0,27;0,40]

Source: European Commission, Joint Research Centre, 2018.

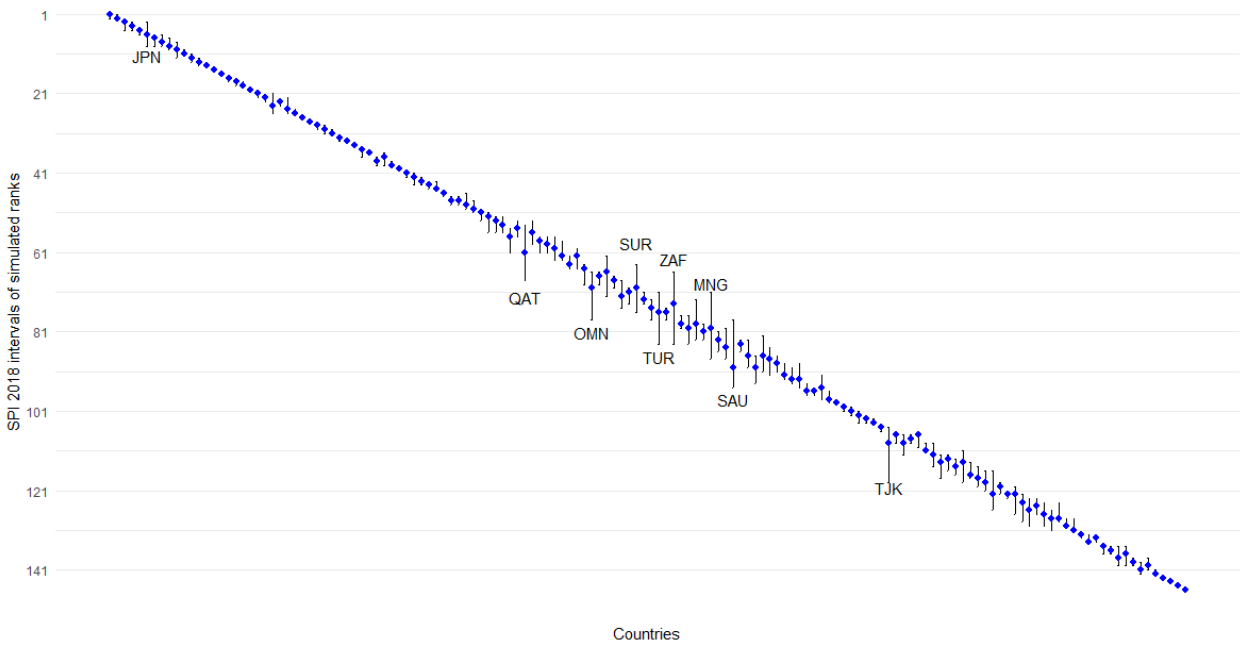
4.1 Uncertainty and Sensitivity analysis

The main results obtained by the robustness analysis are shown in Figure 4, with median ranks and 90% intervals computed across the 2,000 Monte Carlo simulations for SPI. Countries are ordered from best to worst according to their SPI rank, the blue dot represent the median rank among the iterations. Error bars represent, for each country, the 90% interval across all simulations, that is, from the 5th percentile to the 95th. The ranks defined by the SPI always fit in the simulated intervals, and these are narrow enough for most countries (less than 10 positions) to allow for meaningful inferences to be drawn.

SPI ranks are shown to be both representative of a plurality of scenarios and robust to changes in the aggregation method and the dimensions weights. If one considers the median rank across the simulated scenarios as being representative of these scenarios, then the fact that the SPI rank is close to the median rank (less than two positions away) for 85% of the countries suggests that SPI is a suitable summary measure. Furthermore, the reasonable narrow intervals for the majority of the countries' ranks (less than 8 positions) imply that the SPI ranks are robust to changes in the dimensions weights, and the aggregation formula. Only 8 countries are showing a simulated interval that is larger than 10 positions, these countries are labelled in the following figures. Among the top ten ranked countries, Japan (JPN) is the one who has the largest interval (six rank positions).

Overall, country ranks in SPI are fairly robust to changes in the pillar weights and the aggregation formula for the majority of the countries considered. For full transparency and information, Table 5 reports the SPI country ranks together with the simulated intervals (central 90 percentiles observed among the 2,000 scenarios) in order to better appreciate the robustness of these ranks to the computation methodology, and to better analyse the behaviour of specific countries with respect to the perturbations.

Figure 4. Robustness analysis on ranks (SPI rank vs median rank and 90% intervals after perturbation).

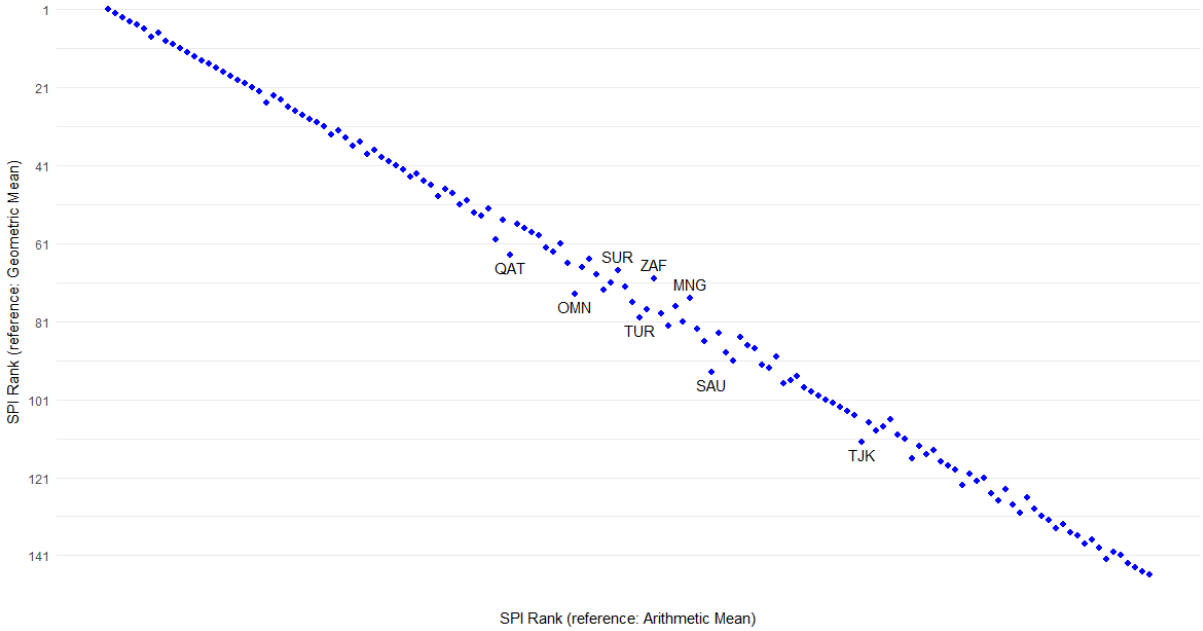


Notes: The countries are ordered respect to their SPI 2018 rank. Median ranks and intervals are calculated over 2,000 simulated scenarios combining simulated weights for the three dimensions (Basic Human Needs, Foundations of Wellbeing and Opportunity) and geometric versus arithmetic average across the three dimensions. The Spearman rank correlation between the median rank and the SPI 2018 rank is 0.9997.

Source: European Commission, Joint Research Centre, 2018.

It is possible to detect only a small group of countries with a larger rank interval (more than 10 positions) in Figure 4. In particular, Tajikistan (TJK), Saudi Arabia (SAU), Oman (OMN) and Qatar (QAT) show low values in the lower border of their intervals, which suggest that some of the scenarios are particularly penalizing their performances. The peculiar behaviour of the countries with large rank intervals can be investigated with sensitivity analysis. In Figure 5, it is possible to compare the ranks derived by SPI with those obtained when changing the aggregation procedure to the geometric mean. This comparison allows investigating if the variability in the rank intervals is originating from the modelling assumption of the aggregation method or by the weights' perturbation. Furthermore, Tajikistan (TJK), Saudi Arabia (SAU), Oman (OMN) and Qatar (QAT) show are positioned below the line, which suggest that their ranks are penalized from a less-compensatory approach such as the geometric mean.

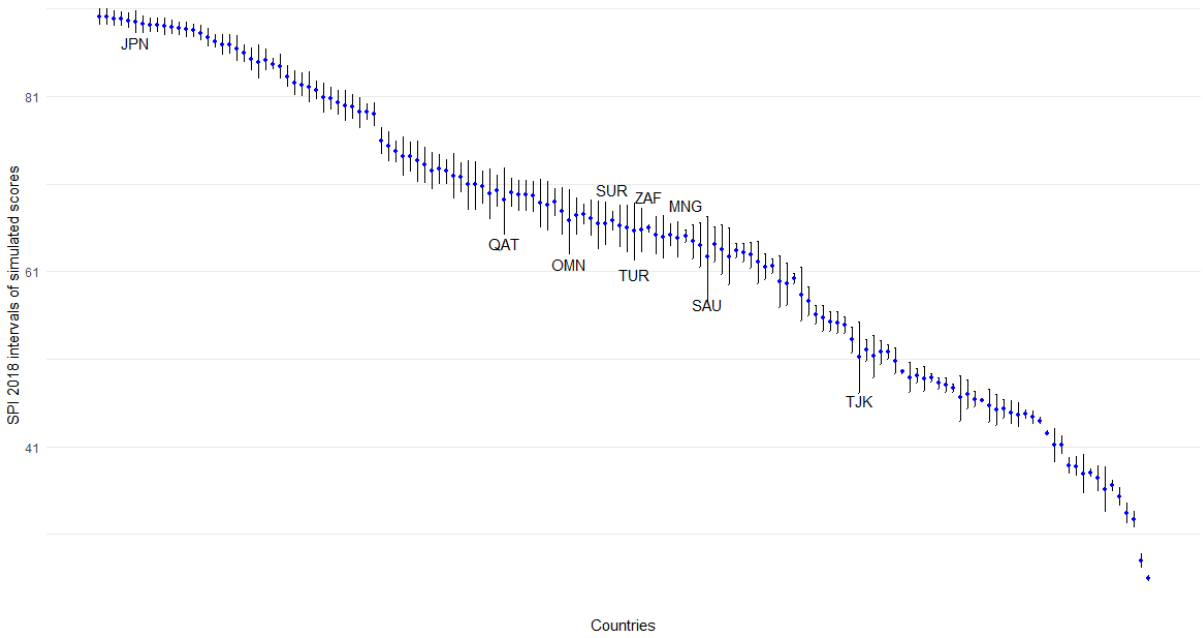
Figure 5. Comparison of ranks according to Arithmetic and Geometric mean.



Notes: The ranks represent the position of the countries in the final index considering the default weights for the three dimensions. The Spearman rank correlation between the two ranks is 0.9987.

Source: European Commission, Joint Research Centre, 2018.

Figure 6. Robustness analysis on scores (median score and 90% intervals of score among perturbations).



Notes: The countries are ordered respect to their SPI 2018 rank. Median scores and intervals are calculated over 2,000 simulated scenarios combining simulated weights for the three dimensions (Basic Human Needs, Foundations of Wellbeing and Opportunity) and geometric versus arithmetic average across the three dimensions. The Spearman rank correlation between the median score and the SPI 2018 rank is 0.9998.

Source: European Commission, Joint Research Centre, 2018.

Some of the variability of ranks can be determined by the perturbation of the weights of the dimensions that influences the resulting score of every country. In the case of SPI, it seems to be a secondary aspect, because most of the difference is determined by the aggregation approach. On the other hand, South Africa (ZAF) and Mongolia (MNG) compared to other countries show very balanced values of the three dimensions and are thus not penalized by the geometric mean. This absence of penalization implies a gain for these two countries' score, which has a large impact on the ranks.

Table 5. SPI rank and 90% interval of all countries.

Country	SPI Rank	Interval	Country	SPI Rank	Interval
Norway	1	[1,2]	Algeria	74	[73,78]
Iceland	2	[1,2]	Turkey	75	[71,84]
Switzerland	3	[3,5]	Morocco	76	[75,78]
Denmark	4	[3,5]	South Africa	77	[66,84]
Finland	5	[5,6]	Kyrgyzstan	78	[77,80]
Japan	6	[3,9]	Bhutan	79	[77,84]
Netherlands	7	[7,9]	Bolivia	80	[73,83]
Luxembourg	8	[6,9]	Dominican Republic	81	[79,83]
Germany	9	[7,10]	Mongolia	82	[71,88]
New Zealand	10	[8,12]	El Salvador	83	[81,86]
Sweden	11	[10,11]	Fiji	84	[80,88]
Ireland	12	[11,13]	Saudi Arabia	85	[78,95]
United Kingdom	13	[12,14]	Paraguay	86	[83,86]
Canada	14	[13,14]	China	87	[83,90]
Australia	15	[15,15]	Iran, Islamic Rep.	88	[87,94]
France	16	[16,16]	Botswana	89	[82,91]
Belgium	17	[17,18]	Philippines	90	[85,92]
Korea, Republic of	18	[17,19]	Indonesia	91	[87,91]
Spain	19	[18,19]	Nicaragua	92	[89,93]
Austria	20	[20,20]	Guyana	93	[90,94]
Italy	21	[21,22]	Sao Tome and Principe	94	[89,95]
Slovenia	22	[22,23]	Egypt	95	[94,97]
Singapore	23	[21,26]	Guatemala	96	[95,97]
Portugal	24	[23,24]	Ghana	97	[92,98]
United States of America	25	[22,26]	Uzbekistan	98	[96,99]
Czech Republic	26	[25,26]	Honduras	99	[98,99]
Estonia	27	[27,27]	India	100	[100,101]
Cyprus	28	[28,28]	Nepal	101	[100,102]
Greece	29	[29,30]	Kenya	102	[101,104]
Israel	30	[29,31]	Senegal	103	[102,104]
Lithuania	31	[30,31]	Timor-Leste	104	[103,104]
Poland	32	[32,33]	Comoros	105	[105,106]
Costa Rica	33	[32,33]	Tajikistan	106	[105,119]
Chile	34	[34,34]	Myanmar	107	[106,109]

Slovakia	35	[35,37]	Bangladesh	108	[107,112]
Hungary	36	[35,36]	Rwanda	109	[107,109]
Croatia	37	[37,39]	Tanzania, United Rep.	110	[106,110]
Uruguay	38	[36,39]	Swaziland	111	[109,111]
Latvia	39	[38,39]	Benin	112	[109,115]
Bulgaria	40	[40,40]	Gambia	113	[112,118]
Barbados	41	[41,42]	Nigeria	114	[112,116]
Argentina	42	[41,44]	Pakistan	115	[113,117]
Mauritius	43	[42,44]	Lesotho	116	[111,119]
Romania	44	[43,45]	Malawi	117	[114,118]
United Arab Emirates	45	[43,45]	Solomon Islands	118	[115,120]
Belarus	46	[46,47]	Togo	119	[116,121]
Serbia	47	[47,49]	Laos	120	[116,126]
Tunisia	48	[47,49]	Cambodia	121	[119,122]
Brazil	49	[46,50]	Côte d'Ivoire	122	[121,123]
Malaysia	50	[48,51]	Liberia	123	[120,127]
Panama	51	[50,53]	Djibouti	124	[122,129]
Montenegro	52	[51,56]	Congo	125	[123,130]
Albania	53	[52,56]	Cameroon	126	[123,127]
Georgia	54	[52,56]	Zimbabwe	127	[124,130]
Armenia	55	[55,61]	Mali	128	[126,131]
Colombia	56	[53,57]	Sierra Leone	129	[124,129]
Qatar	57	[54,68]	Mozambique	130	[128,130]
Mexico	58	[53,59]	Burkina Faso	131	[128,131]
Ecuador	59	[57,61]	Madagascar	132	[132,133]
Russian Federation	60	[57,61]	Mauritania	133	[132,134]
Peru	61	[57,63]	Ethiopia	134	[133,134]
Jordan	62	[58,63]	Angola	135	[135,137]
Cuba	63	[62,65]	Guinea	136	[135,137]
Ukraine	64	[60,65]	Sudan	137	[135,140]
Macedonia, FYR	65	[64,69]	Papua New Guinea	138	[135,140]
Oman	66	[66,78]	Burundi	139	[138,140]
Sri Lanka	67	[66,69]	Yemen	140	[139,142]
Cabo Verde	68	[62,72]	Niger	141	[138,141]
Moldova, Republic of	69	[67,70]	Congo, Dem. Rep.	142	[141,142]
Thailand	70	[68,75]	Eritrea	143	[143,143]
Kazakhstan	71	[70,74]	Afghanistan	144	[144,144]
Suriname	72	[64,76]	Chad	145	[145,145]
Lebanon	73	[71,74]	Central African Republic	146	[146,146]

Notes: The countries are ordered respect to their SPI 2018 rank. The intervals are calculated over 2,000 simulated scenarios combining simulated weights for the three dimensions (Basic Human Needs, Foundations of Wellbeing and Opportunity) and geometric versus arithmetic average across the three dimensions.

Source: European Commission, Joint Research Centre, 2018.

5 Basic Human Needs, Foundations of Wellbeing and Opportunity - three similar and complementary concepts of Social Progress

The added value of the Social Progress Index as a summary measure of the three dimensions will be discussed in this section and how the statistical associations between the three dimensions can be used to inform social policies at national level.

Table 6 shows that the SPI ranking and any of the three dimension rankings differ by 10 positions or more for at least 1/5 of countries. This finding suggests that there is an added value in referring to the SPI results in order to identify aspects of a country's social progress that do not directly emerge by looking into the three dimensions separately. At the same time, this outcome points to the value of examining individual dimensions on their own merit in order to see which aspects of social progress are driving a country's performance.

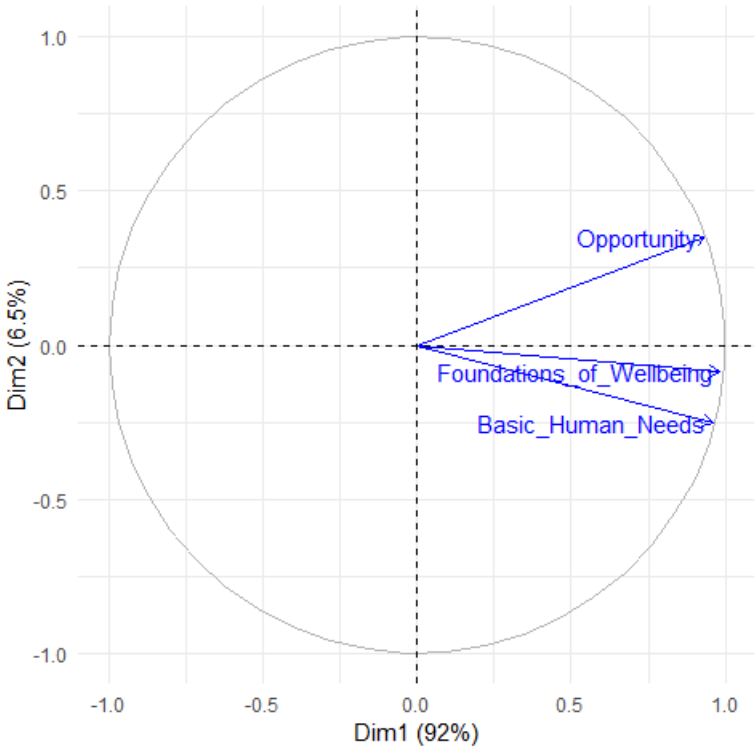
Table 6. Distribution of differences between dimension and ESI rankings.

Shifts with the respect to SPI	Basic Human Needs Dimension	Foundations of Wellbeing Dimension	Opportunity Dimension
0 positions	10%	4%	6%
Less than 5 positions	36%	46%	43%
5 to 9 positions	30%	33%	18%
More than 10 positions	**34%**	**21%**	**39%**
10 to 19 positions	25%	20%	22%
20 to 29 positions	8%	1%	8%
More than 30 positions	1%	0%	9%
Total	100%	100%	100%

Source: European Commission, Joint Research Centre, 2018.

As already pointed out in Section 3, PCA was used to confirm the existence of one single statistical dimension among the three SPI dimensions. The first principal component explains about 92% of the total variation, and the second principal component an additional 6,5%. Thus the first two principal components explain nearly all variance in the data (99% of the total variance). Figure 7 illustrates the projections of the SPI dimensions onto the plane spanned by the first two principal components in a so called "factor map".

Figure 7. Factor map of the three SPI dimensions Basic Human Needs, Foundations of Wellbeing and Opportunity.



Source: European Commission, Joint Research Centre, 2018.

The correlation between each SPI dimension and the principal component is given by the projection of the SPI dimension vector onto the component axis. The three SPI dimensions are all correlating highly with the first principal component. The correlations are in the range from 0,93 (third dimension Opportunity) to 0,98 (second dimension Foundations of Wellbeing). The second principal component is much less influential than the first and is only accounting for less than 1/15 of the total variance. Despite being less influential, the second principal component is useful to evaluate the differences between the first two and third dimensions. This difference is illustrated in Figure 7, where the first two dimensions point to the lower part of the graph remaining very close, and the third dimension points up showing a difference with respect to the others. The similarities between dimensions are further investigated in the bivariate scatter plots of the SPI dimensions in Figure 8. For the first two dimensions, Basic Human Needs and Foundations of Wellbeing, there is an excellent linear relationship with a correlation coefficient of 0,95. The linear relationships are very good but not as strong for the other two combinations. The correlation coefficient is 0,81 for the Basic Human Needs and Opportunity dimensions and 0,88 for the Foundations of Wellbeing and Opportunity. Given the complexity of the intrinsic SPI structure and the variety of countries at all levels of economic development, the interpretation of the associations between the dimensions require some special attention. Figure 8(a) depicts the relation between the first two dimensions Basic Human Needs and Foundations of Wellbeing. The identity line (black line) enables an easy understanding of the comparison. Most countries show a higher value of Basic Human Needs respect to Foundations of Wellbeing. However 14 countries among the lower half of Social Progress have achieved a better Foundations of Wellbeing score. These countries are all situated in Africa; Zimbabwe (ZWE) has the largest score difference. For highly developed countries there is little variability between

the two dimensions, but below the score 85 the dimensions are more variable. For instance, Uzbekistan (UZB) and Zimbabwe (ZWE) share the same score of Foundation of Wellbeing (both scoring 56), but their results on Basic Human Needs differ largely (respectively 82 and 44). Many other comparisons of this kind may be found, which proves the difference between the two dimensions, and justify their contextual existence in the index.

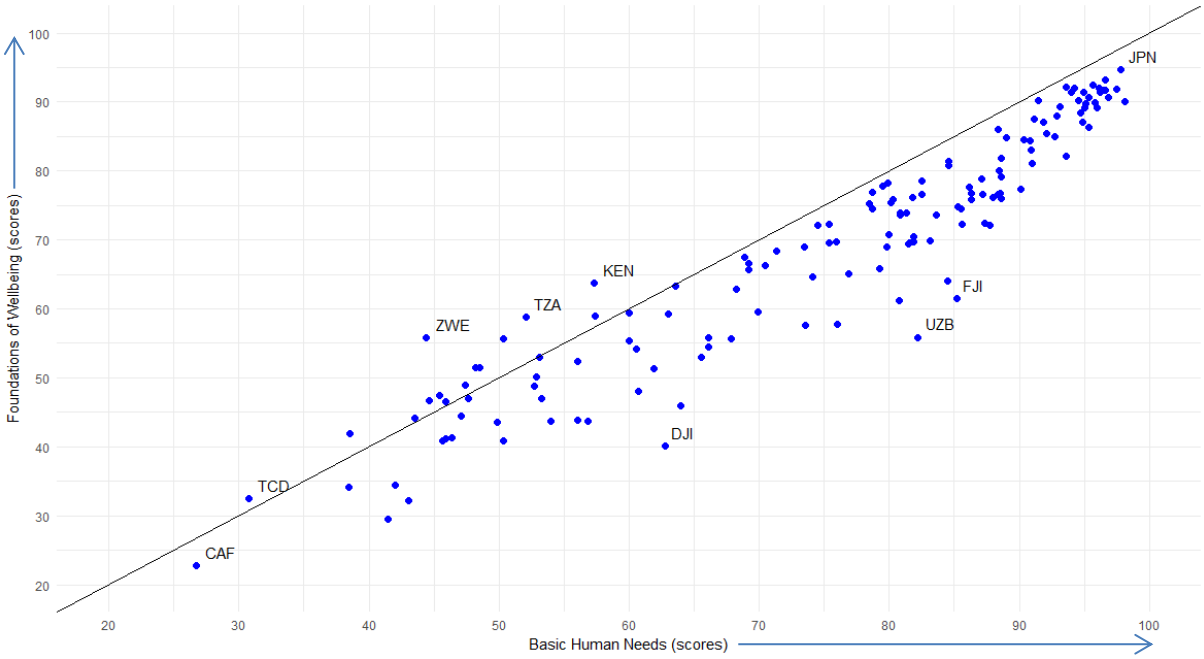
The relationship between the dimensions of Basic Human Needs and the Opportunity is represented in the second graph Figure 8(b). The Opportunity scores are generally lower than the Basic Human Needs scores. This result is not surprising, because it is possible to assume the third dimension as a successive effect of a decent level in the first two. It is difficult to assume a satisfying level of Opportunity without a decent level of Basic Human Needs and Foundations of Wellbeing. The distribution of the countries in this comparison is even more informative than the previous one. Apart from the top ranked countries, the two dimensions are representing very diverse scores. Saudi Arabia (SAU) and Tajikistan (TJK), for instance, show low level of Opportunity while performing high values in the Basic Human Needs dimension. Another informative example is Yemen (YEM) and Lesotho (LSO), they both show low scores of Basic Human Needs (respectively 50 and 49), but are radically different from the perspective of Opportunities (respectively 21 and 54).

The analysis of the last graph in Figure 8(c) can easily be read in the light of the previous ones. For example, Uruguay (URY) and Qatar (QAT) are both performing well in the Foundations of Wellbeing dimension but have very diverse scores in the Opportunity dimension. For most countries, the values of Foundations of Wellbeing, as well as for Basic Human Needs, are higher than for the values of Opportunity. This remains true except for a few African countries ranked at the lower end of Social Progress.

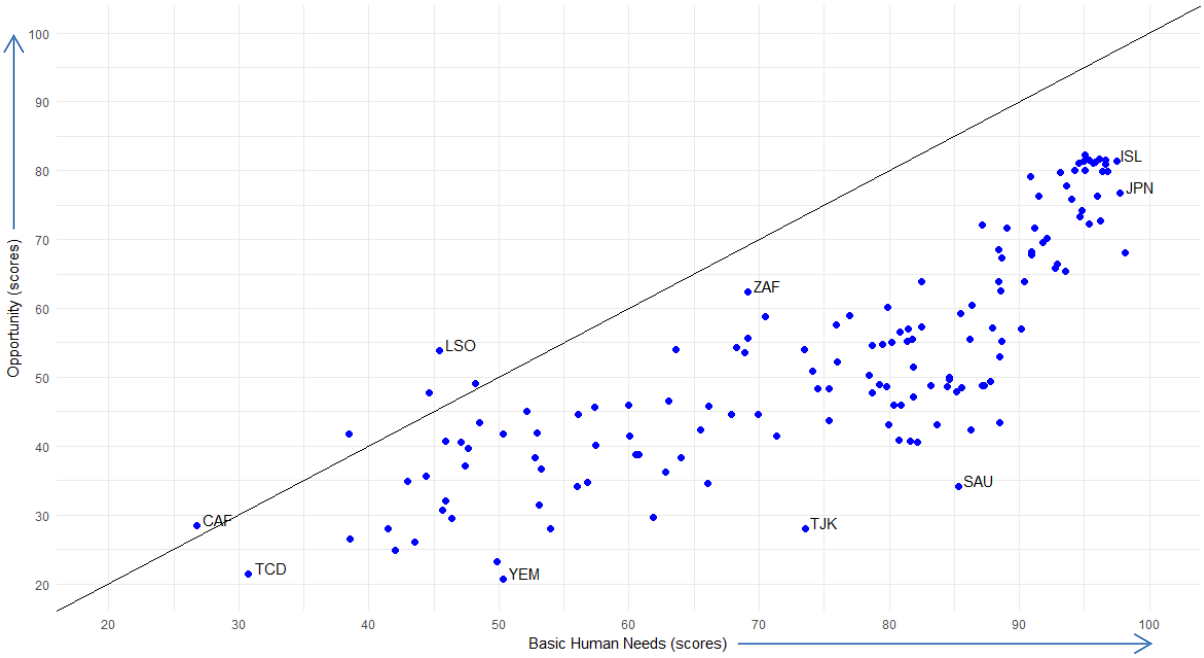
The three dimensions are showing different aspects and specificities of Social Progress and they are all giving a contribution to the informative and enlightening power of the Social Progress Index.

Figure 8. Scatter plots of SPI 2018 Dimensions.

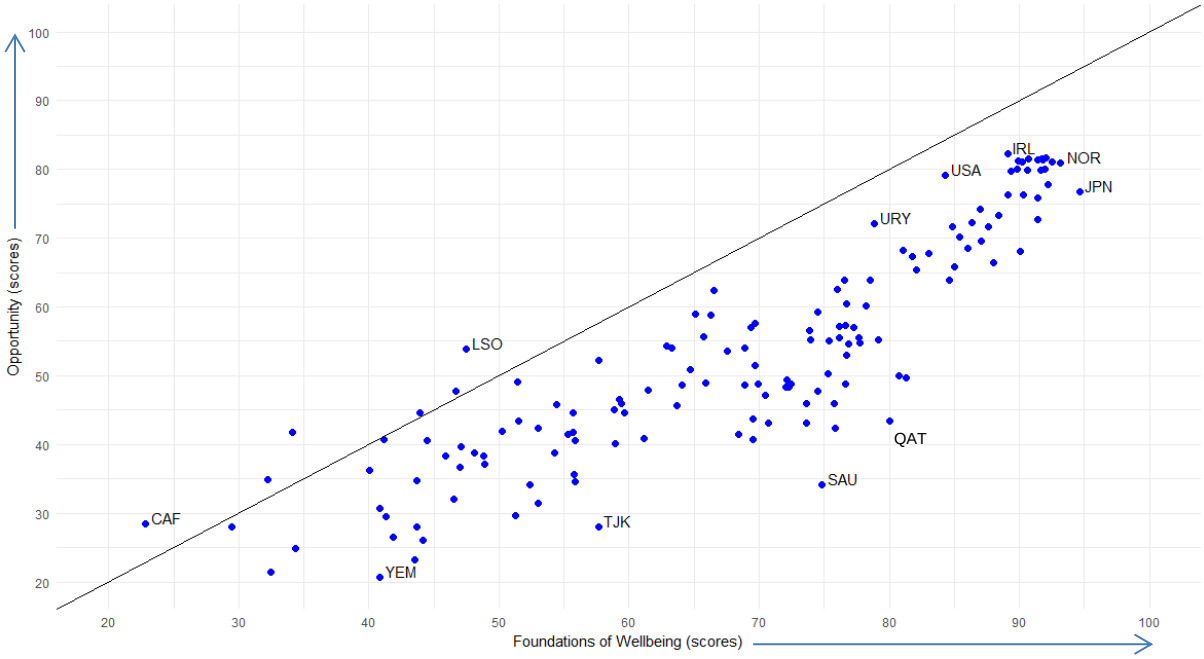
a. Dimension 1 Basic Human Needs vs Dimension 2 Foundations of Wellbeing.



b. Dimension 1 Basic Human Needs vs Dimension 3 Opportunity.



c. Dimension 2 Foundations of Wellbeing vs Dimension 3 Opportunity.



Notes: The dots represent country scores for the three SPI dimensions: Basic Human Needs, Foundations of Wellbeing and Opportunity for the 146 ranked countries. The straight line is the line of equality (identity line).

Source: European Commission, Joint Research Centre, 2018.

6 Conclusions

The Social Progress Imperative has developed the Social Progress Index (SPI) which is an international monitoring framework for measuring social progress without resorting to the use of economic indicators. It provides a basis to understand the relationship between economic and social progress and measures country performance on aspects of social and environmental performance. The index ranks 146 countries, from less developed to well-developed, on social progress. The JRC statistical audit has delved around in the workings of the SPI framework to assess the statistical properties of the data, and the methodology used in the index construction.

The SPI framework is well-constructed, and a lot of research and thought have been devoted in its development. The SPI is hierarchically structured, based on three dimensions: Basic Human Needs, Foundations of Wellbeing and Opportunity, twelve components and 51 social outcome indicators.

The key findings of the statistical assessment conducted herein are the following:

First, the statistical coherence analysis suggests that the three-level structure in the SPI framework and the "unidimensionality" of all SPI variables (components, dimensions and index) are confirmed. Furthermore, all 51 indicators were found to be influential from the first aggregation level (component level) up to the index level. This means that the information content in all underlying indicators is maintained at all levels of aggregation in the SPI framework.

Second, the statistical coherence analysis has also pointed out that three components are fitting better in other dimensions compared to where they have been allocated. The three components are the Personal Safety (C4), the Personal Freedom and Choice (C10) and Access to Advanced Education (C12).

Third, some redundancies may appear in the SPI framework due to high correlations between the different SPI aggregates. Indeed, PCA has shown that the number of components in the SPI framework may be reduced. Seven to nine components would suffice to explain social progress adequately. Fewer components are in line with the results from the "beyond GDP" approach by Stiglitz *et al.* which identifies eight key dimensions of well-being. A simpler SPI framework with less components and/or indicators would also facilitate data collection and descriptive power of the index.

Forth, the three dimensions of the SPI: Basic Human Needs, Foundations of Wellbeing and Opportunity, may at first glance seem to be measuring very similar concepts according to their pairwise correlations. They are however complementing each other especially for the lower ranked countries. Every dimension induces a different ranking respect to the SPI, and the pairwise comparisons of the values of the dimensions show how they introduce additional information about the countries.

Fifth, the choice of arithmetic mean for the aggregation is influential especially for mid ranked countries because of their diverse score profiles. For the top ranked countries, on the other side, the aggregation procedure is not so influential because their scores are usually more balanced among the dimensions. The arithmetic mean is a central feature in the index construction which has to be taken into account in its utilization, in order to make correct interpretations of the results.

Sixth, the ranks resulting from the uncertainty analysis prove a satisfying independence of the index from the choice of weights and aggregation procedure. Despite the robustness of the index, there are some countries showing rank variability. This is an expected feature and proves there is room for change of the rankings. The index appears robust and easy to use for the informative aims it was made for.

The present JRC audit findings confirm that the Social Progress Index 2018 appears to be a comprehensive quantitative method to measure and monitor social progress at national level worldwide. It is a conceptually and statistically sound tool that is widely applicable for ongoing assessment of social progress and a potential benchmark against which to compare future progress. Focusing on the different constituents of the SPI, country-specific aspects of social progress can be identified and serve as an input for data-informed policy analysis on social progress.

7 References and related reading

Becker, W., M. Saisana, P. Paruolo, and I. Vandecasteele. 2017. 'Weights and Importance in Composite Indicators: Closing the Gap'. *Ecological Indicators* 80: 12–22.

Groeneveld, R. A. and G. Meeden. 1984. 'Measuring Skewness and Kurtosis'. *The Statistician* 33: 391–99.

Little, R. J. A. and D. B. Rubin. 2002. *Statistical Analysis with Missing Data*. 2nd edition. Hoboken, NJ: John Wiley & Sons, Inc.

Munda, G. 2008. *Social Multi-Criteria Evaluation for a Sustainable Economy.* Berlin Heidelberg: Springer-Verlag.

OECD/EC JRC (Organisation for Economic Co-operation and Development/European Commission, Joint Research Centre). 2008. *Handbook on Constructing Composite Indicators: Methodology and User Guide.* Paris: OECD.

Paruolo, P., M. Saisana, and A. Saltelli. 2013. 'Ratings and Rankings: Voodoo or Science?' *Journal of the Royal Statistical Society* A 176 (3): 609–34.

Saisana, M., B. D'Hombres, and A. Saltelli. 2011. 'Rickety Numbers: Volatility of University Rankings and Policy Implications'. *Research Policy* 40: 165–77.

Saisana, M., A. Saltelli, and S. Tarantola. 2005. 'Uncertainty and Sensitivity Analysis Techniques as Tools for the Analysis and Validation of Composite Indicators'. *Journal of the Royal Statistical Society* A 168 (2): 307–23.

Saltelli, A., M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. 2008. *Global Sensitivity Analysis: The Primer*. Chichester, England: John Wiley & Sons.

8 List of figures and tables

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

- at the following standard number: +32 22999696, or

- by electronic mail via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications
You can download or order free and priced EU publications from EU Bookshop at: https://publications.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

KJ-NA-29576-EN-N

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

Joint Research Centre

EU Science Hub

Publications Office

doi:10.2760/343522

ISBN 978-92-79-98411-2