JRC Statistical Audit of Commitment to Reducing Inequality Index

Michaela Saisana and Béatrice d'Hombres

European Commission, Joint Research Centre (JRC), Competence Centre on Composite Indicators and Scoreboards (COIN)

Conceptual and practical challenges are inevitable when trying to summarise with a single composite indicator the commitment of countries to reducing inequality. The summary and section 5 of the Commitment to Reducing Inequality Index report discusses in detail the conceptual framework, while the selection of indicators, data quality aspects and methodological choices for grouping country-level data across 8 main indicators, 3 pillars and an overall index are presented in the Methodology Annex.

The statistical audit presented below constitutes the first collaboration between Oxfam and the European Commission's Joint Research Centre (JRC), specifically the Competence Centre on Composite Indicators and Scoreboards. The statistical assessment carried out by JRC aims to contribute to ensuring the transparency and reliability of the Commitment to Reducing Inequality (CRI) index and thus to enable policymakers to derive more accurate and meaningful conclusions, and to potentially guide choices on priority setting and policy formulation.

Statistical soundness should be regarded as a necessary but not a sufficient condition for a sound index, since the correlations underpinning the majority of the statistical analyses carried out herein "need not necessarily represent the real influence of the individual indicators on the phenomenon being measured".[1] The development of any index must thus be nurtured by a dynamic iterative dialogue between the principles of statistical and conceptual soundness. In that respect, prior to undertaking the present statistical assessment, Oxfam and JRC engaged in previous discussions during spring 2017. An earlier version of the CRI index was assessed by the JRC in March–April 2017. Fine-tuning suggestions, aimed at setting the foundation for a balanced index, were taken into account by Oxfam and Development Finance International research teams for the final computation of the CRI scores and rankings.

The JRC assessment of the CRI index presented in this appendix has focused on two main issues: the statistical coherence of the structure, and the impact of key modelling assumptions on the CRI scores and ranks.[2] In particular, the JRC analysis complements the reported country rankings for the CRI index with estimated confidence intervals, in order to better appreciate the robustness of these

[1] OECD & JRC (2008)

[2] The JRC analysis was based on the recommendations of the OECD & JRC (2008) Handbook on Composite Indicators, and on more recent research from the JRC. The JRC auditing studies of composite indicators are available at https://ec.europa.eu/jrc/en/coin (all audits were conducted upon request of the Index developers).

ranks to some modelling choices (such as the weighting scheme, aggregation formula and estimation of missing values).

Overall, the main conclusions of the present audit can be summarised as follows: the CRI Index is representative of a plurality of scenarios, reliable and with a statistically coherent framework. The uncertainty analysis shows that country ranks are robust for one-third of the countries. For a number of countries, in particular non-OECD countries, ranks should be analysed within their expected confidence intervals instead of being taken at face value. The statistical assessment has also shown that the CRI index has a good statistical reliability and measures one single latent phenomenon capturing the three main components of the index, the Spending, Tax and Labour pillars. Notwithstanding the good statistical properties of the CRI index, some suggestions are made for possible refinements of the CRI index in future editions once more data and time series are available.

1. Statistical Coherence in the CRI Framework

The twenty-one data points used for each country to derive the eight main indicators that compose the overall CRI index have been selected by Oxfam and Development Finance International research teams for their relevance with the conceptual framework, on the basis of the literature review, expert opinion and timeliness. The conceptual relevance of the indicators underpinning the CRI framework is thus not discussed in this appendix. The assessment of the statistical coherence of the CRI Index starts from the level of the eight main indicators grouped across the three pillars and further aggregated into an overall CRI index.

The present statistical assessment of the CRI index has involved the following steps:

1.1 Data Checks

Data coverage is very good. The three indicators composing the pillar ‘Progressivity of Spending’ do not include any missing values in the final dataset. For the second pillar ‘Progressivity of Tax’, only the indicator measuring the Tax collection effort is affected by a few missing values (9.2% of the countries). Finally, the three indicators included in the third pillar ‘Progressivity of Labour Policies’ have missing values for 8.5% of countries. This implies that for those countries the CRI overall scores are based on the performance on the first two pillars. In the uncertainty analysis, the robustness of CRI ranking to an alternative imputation method will be tested.

Potentially problematic indicators that could bias the overall index were considered as those having absolute skewness greater than 2 and kurtosis higher than 3.5. As shown in Table 1, none of the 8 indicators used in the calculation of the CRI Index are highly skewed.

Table 1: Summary Statistics

CRI components	Number of Observations	Missing data (%)	Mean	skewness	kurtosis	[Min,Max]
Pillar 1 : Progressivity of Spending	152	0	0.378	0.958	-0.265	[0,1]
Government spending on Progressive Sectors	152	0	0.368	0.141	-0.63	[0.05, 0.76]
Incidence of Spending	152	0	0.252	1.397	0.539	0,1]
Pillar 2 : Progressivity of Tax -	152	0	0.523	-0.21	-0.355	[0,1]
Tax Structure	152	0	0.585	-0.688	0.276	[0.15,0.92]
Tax Incidence	152	0	0.677	-1.127	2.897	[0,1]
Tax Collection	138	9.2	0.425	0.399	0.069	[0,1]
Pillar 3: Progressivity of labour Policies	139	8.5	0.431	0.5	-0.709	[0,1]
Labour Union Rights	139	8.5	0.482	0.17	-0.304	[0,1]
Women's rights in the Workplace	139	8.5	0.338	0.411	-1.089	[0,1]
Minimum Wage as a % GDP	139	8.5	0.313	1.074	1.463	[0,1]

Source: European Commission, Joint Research Centre (JRC), 2017

Notes: Rule for outlier detection: Indicators with a |skewness|>2 and a kurtosis >3.5

1.2 Statistical Coherence

The statistical coherence consists of a principal components analysis to explore the structure of the data, a multi-level analysis of the correlations of variables, and a comparison of CRI rankings with its pillars and with other indicators measuring the actual level of inequality.

Principal components analysis and cross-correlation analysis

Principal component analysis (PCA) has been used to assess the extent to which the conceptual framework is compatible with the statistical properties of the data. The PCA has been carried out at the pillar level.

Ideally, PCA should confirm the presence of a single statistical dimension amongst the indicators subject to analysis, i.e., no more than one principal component with eigenvalue greater than 1.0. The PCA suggests that the three pillars share a single latent dimension that summarises 63% of the total variance, with the correlation coefficients of the three pillars with the first principal component ranging between 0.46 and 0.64. The first Pillar on 'Progressivity of Spending' and the third Pillar measuring the 'Progressivity of Labour Policies' contribute equally to the first principal component, while the second 'Pillar on the Tax Structure' weights comparatively less. The reliability of the aggregate of the three dimensions is satisfactory with a Cronbach-alpha value equal to 0.71.[3]

An earlier version of the CRI framework that was analysed by the JRC during March–April 2017 included a ninth indicator on 'Tax haven status'. Preliminary PCA results showed however that the

[3] The Cronbach's alpha is a measure of correlations for all pairs of indicators which assesses the reliability of the indicators composing the dimension. When the Cronbach's alpha is above 0.7, the indicators are considered to reliably measure the underlying dimension.

data for this area of policy was not consistent with the data used for the rest of the indicators in the CRI framework and it was hence recommended not to incorporate it into the Index itself and to use it instead as a supplementary data point.

Table 2: Pairwise correlations between indicators, pillars and the CRI index

	Spending pillar	Tax pillar	Labour pillar	CRI Index
Government spending on Progressive Sectors	0.83	0.25	0.45	0.71
Incidence of Spending	0.96	0.35	0.69	0.71
Tax Structure	-0.22	0.53	n.s	n.s
Tax Incidence	0.29	0.56	0.19	0.42
Tax Collection	0.46	0.47	0.37	0.51
Labour Union Rights	0.68	0.29	0.81	0.76
Women's rights in the Workplace	0.71	0.29	0.84	0.79
Minimum Wage as a % GDP	n.s	n.s	0.41	n.s

Source: European Commission, Joint Research Centre (JRC), 2017.
Notes: Numbers represent Pearson correlation coefficients. 'n.s': Non-significant correlations at the 99% level.

A more detailed analysis of the correlation structure within and across the three pillars confirms the expectation that the indicators are more strongly associated with their own pillar than to any of the other two pillars (see Table 2). This suggests that the allocation of the indicators to the specific pillar is consistent both from conceptual and statistical perspectives.

The correlations between the indicators and the CRI Index show that six out of the eight indicators are positively and significantly correlated with the overall CRI Index (Table 2). Yet two indicators – 'Tax structure' and 'Minimum Wage as a % GDP'– are not influential at the overall CRI level though they are at their own pillar level. This is indicative of a different behaviour of these two indicators with respect to the remaining ones composing the CRI index.

At the overall CRI level, the three pillars correlate strongly with the CRI index, with pairwise correlations above 0.60 (see Table 3). The Spending and Labour pillars are, however, more strongly associated with the overall CRI index than the Tax pillar. Had the re-normalization not been applied at the pillar level, the result would have been more polarised and the correlation of the Tax pillar to the CRI Index would have been close to 0.5, while the correlation between the Spending pillar and the CRI Index would have been at 0.94.

Table 3: Pairwise correlations between pillars and the CRI index

	Spending pillar	Tax pillar	CRI Index
Spending pillar	1	0.34	0.89
Tax pillar	0.34	1	0.61
Labour pillar	0.66	0.29	0.84

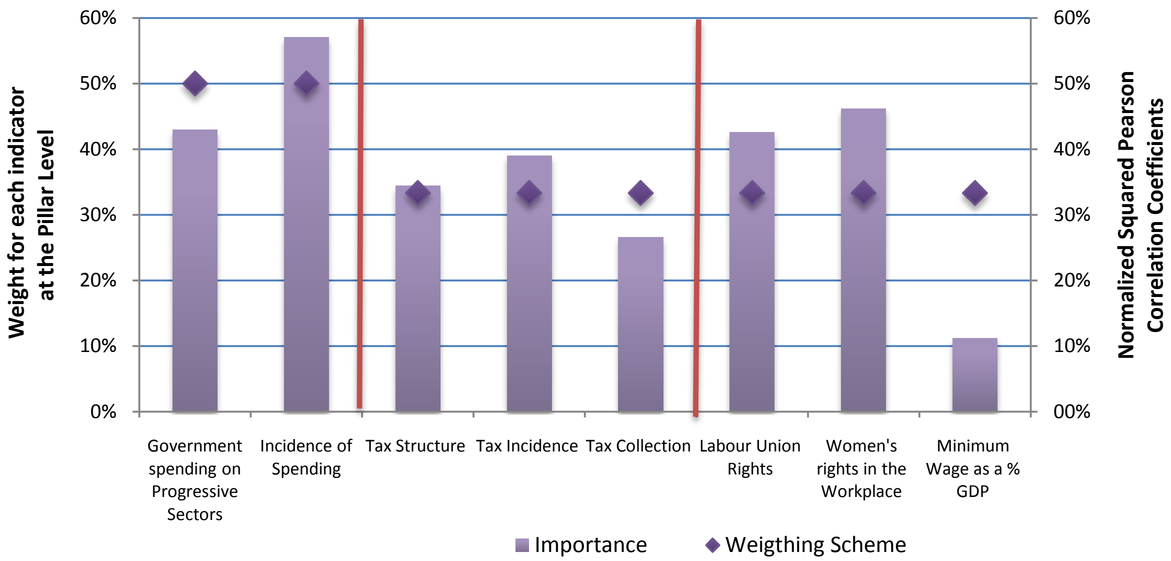
Source: European Commission, Joint Research Centre (JRC), 2017.
Notes: Numbers represent Pearson correlation coefficients.

Importance of the indicators in the CRI framework

Each of the three pillars composing the CRI Index is a simple arithmetic average of the underlying indicators. Developers and users of composite indicators often consider that the weights assigned to the indicators coincide with the indicators' importance in the index. However, in practice, the correlation structure of the indicators and their different variances do not always allow the weights assigned to the indicators to be considered equivalent to their importance.

This section assesses the importance of all eight indicators at various levels of aggregation in the CRI structure. As a statistical measure of the importance of indicators we rely on the normalised squared Pearson correlation coefficients. The squared Pearson correlation coefficient measures the percentage of the variance of the pillar (or CRI Index) scores that is explained by each indicator. The result of our analysis comparing the actual importance of the indicators with their original weight is reported in Figure 1. The dots correspond to the weights assigned to each indicator within the pillar to which the indicator belongs, while the bars represent the actual statistical importance of the indicators.

Figure 1: Weights and Statistical Importance of the indicators within each pillar



Source: European Commission, Joint Research Centre (JRC), 2017.

Within the first pillar, the indicator on the incidence of spending captures a larger proportion of the variance of the Spending pillar compared with the indicator measuring the government spending on progressive sectors (57% versus 43%). Given that an equal weight was applied to the two indicators, this implies that the weighting scheme slightly deviates from the actual importance of the indicators. For the second pillar on the Progressivity of Tax, the three indicators have a relatively similar importance with the normalized Pearson correlation coefficients ranging between 27% and 39%. The most noteworthy difference between the equal weight assigned to the indicators and their actual statistical importance is found for the third pillar on the Progressivity of Labour Policies. Indeed, the first two indicators on the 'Respect for labour and union rights' and 'Legal protection for women in the work place' are much more influential than the third indicator measuring the fairness of the level of minimum wages. This result is in line with the correlation analysis results and suggests that the CRI developing team might need to reconsider how best to include this indicator in the next year's release, should this statistical result be confirmed also with next year's dataset. For this first release, it is only flagged for attention.

1.3 Added value of the CRI Index

A high statistical reliability among the main components of an index can be the result of redundancy of information. This is clearly not the case with the CRI index. For 60% or more of the 152 countries included in the CRI Index, the CRI ranking and any of the three pillar rankings differ by 10 positions or more (Table 4). This suggests that the CRI ranking highlights aspects of countries' efforts to reducing inequality that do not emerge by looking into the three pillars separately. At the same time, this result points to the value of examining individual pillars and indicators on their own merit.

Table 4: Distribution of differences between pillars and CRI rankings

Shifts with respect to CRI index	Spending pillar	Taxation pillar	Labour pillar
More than 30 positions	24%	44%	29%
20 to 29 positions	11%	9%	15%
10 to 19 positions	26%	20%	24%
5 to 9 positions	23%	11%	17%
Less than 5 positions	14%	14%	15%
0 positions	3%	1%	0%
Total	100%	100%	100%
More than 10	60%	74%	62%

Source: European Commission, Joint Research Centre (JRC), 2017.

The CRI index was also compared with country rankings based on actual levels of inequality. The two measures of inequality used for this purpose are the Gini Coefficient and the Palma index.[4] The Palma index is the ratio of the income share of the top 10% to that of the bottom 40%. The correlations between these two measures of inequality and the CRI Index are low on the full-sample with pairwise correlation coefficients respectively equal to -.29 and -.23 with the GINI and Palma measures, while these two numbers amount to -.66 and -.67 when the statistics are computed on the sub-sample of OECD countries (Table 5). Along the same line, Table 5 shows countries in the 25th percentile in terms of both inequality measures have an average CRI score substantially higher than those situated in the 75th percentile, this being particularly true for OECD countries.

Though we should not interpret this as a causal relationship, this suggests that OECD countries with low levels of inequalities are those putting more efforts to ensure a more equity-based society.

Table 5: Comparison of the CRI Index with other inequality-based rankings

	CRI index	
Gini coefficient and Palma index	Full sample	OECD countries
• Countries in the 75th percentile	0.36	0.54
• Countries in the 25th percentile	0.53	0.79
Correlation coefficients		
• Gini coefficient	-0.29	-0.66
• Palma index	-0.22	-0.67

Source: European Commission, Joint Research Centre (JRC), 2017.

[4] We used the two most recent years available (World Development Indicators database)

2. Impact of modelling assumptions on the CRI results

Country scores depend both on the data underlying the selected indicators and on modelling choices. The three-pillar structure and choice of indicators, the treatment of missing data, the normalization method, weighting scheme, and aggregation formula, are, among other elements, all impacting on CRI-based country ranks. These choices are based on expert opinion (e.g. selection of indicators), common practice (e.g. min-max normalization in the [0, 1] range for both indicators and pillars) or simplicity (e.g. no imputation of missing data).

The robustness analysis in this section aims at assessing the simultaneous and joint impact of these modelling choices on the CRI rankings. The data underpinning the indicators are assumed to be error-free, since potential outliers and eventual errors and typos were corrected during the computation phase by Oxfam and Development Finance International research team.

The robustness assessment of the CRI is based on a combination of a Monte Carlo experiment and a multi-modelling approach as it is commonly done in the relevant literature on composite indicators (Saisana *et al.*, 2005; Saisana *et al*., 2011). Three methodological assumptions have been included in the uncertainty analysis: (a) the choice of not estimating missing values, (b) the weight assigned to each pillar, and (c) the aggregation formula used to compute the CRI overall score.[5] This type of uncertainty analysis aims to complement CRI country ranks with confidence intervals in order to help users of the index to appreciate for which countries ranks can be taken at face value and for which countries instead country ranks are to be analysed with caution because of their sensitivity to the methodological choices underlying the index computation.

The Monte Carlo simulations relate to the issue of weighting, and comprised 1,000 runs. Each run corresponds to a different set of weights assigned to each of the three pillars. The weights were randomly sampled from uniform continuous distributions centered at the weight value originally adopted for calculation of the CRI score (=1/3). A perturbation of the weights ± 25% around the reference values was adopted. For each simulation, weights are rescaled so that they always sum up to 1. The choice of the range for the weights' variation was driven by two opposite needs: ensure a wide enough interval to have meaningful robustness checks; and respect the rationale of the CRI that places on an equal footing all three pillars. Given these considerations, limit values of uncertainty intervals for the pillar weights are 25–42% for each pillar (see Table 6).

For reasons of transparency and replicability, the CRI developing team opted not to estimate the few missing values (see Table1). The 'no imputation' choice – common in similar contexts of index development – might encourage countries not to report low data values. As mentioned earlier, missing values in the CRI framework are primarily concentrated in the Labour pillar and exclusively in

[5] Note that other uncertain parameters entering into the calculation of the CRI score could have been taken into account. However, previous uncertainty analyses have shown that these three assumptions (aggregation method, weighting scheme and imputation methods) are those having the strongest impact on composite indicators-based rankings.

non-OECD countries.[6] To test the impact of this assumption, the JRC estimated missing values using the Expectation Maximization (EM) algorithm.[7]

Regarding the aggregation formula, decision-theory practitioners challenge the use of simple arithmetic averages because of their fully compensatory nature, in which a comparative high advantage on a few indicators can compensate a comparative disadvantage on many indicators (Munda, 2008). These challenges are known to the CRI developers who opted to adopt a mixed aggregation formula that was chosen with the following rationale: "… *while both tax and spending can be individually progressive, a greater commitment to reducing inequality is demonstrated when both tax and spending act together*". To capture this interaction, the developers have multiplied the spending score by the tax score. Furthermore, it was assumed that spending, tax and labour market policies are equally important to a country's commitment to reducing inequality. With a view to place the three pillars – tax, spending, labour market – on equal footing while accounting for the interaction between tax and spending policies, each country's CRI score is computed as an arithmetic average made up by two-thirds of the geometric average of the tax and spending pillars while one-third is assigned to the labour pillar. In order to test for the impact of this aggregation formula at the pillar level, the JRC considered as alternative the geometric average between all three pillars. This aggregation method is a partially compensatory approach that rewards countries with similar performance in all pillars, and motivates those countries with uneven performance to improve in those pillars in which they perform poorly, and not just in *any* pillar.[8]

Table 6: Uncertainty analysis for the CRI: weights, missing data, normalization, aggregation

	Reference	Alternative
I. Uncertainty in the treatment of missing values (indicator level)	No estimation of missing data	Expectation Maximization (EM)
II. Uncertainty intervals for the CRI weights (pillar level)	Reference value for the weight	Distribution assigned for robustness analysis
Spending Pillar (S)	0.33	U[0.25,0.42]
Taxation Pillar (T)	0.33	U[0.25,0.42]
Labour Rights Pillar (L)	0.33	U[0.25,0.42]
III. Uncertainty in the aggregation formula (pillar level)	$=\frac{\sqrt{S\times T}+\sqrt{S\times T}+L}{3}$	Geometric average

Source: European Commission, Joint Research Centre (JRC), 2017.

[6] The 'no imputation' choice for missing values implies a 'shadow imputation'. With arithmetic averages the absence of imputation is equivalent to replacing missing values with the average of the available (normalized) scores.

[7] The Expectation-Maximization (EM) algorithm (Little and Rubin, 2002) is an iterative procedure that finds the maximum likelihood estimates of missing values by repeating two steps: (1) The expectation E-step: Given a set of parameter estimates, such as a mean vector and covariance matrix for a multivariate normal distribution, the E-step calculates the conditional expectation of the complete-data log likelihood given the observed data and the parameter estimates. (2) The maximization M-step: Given a complete-data log likelihood, the M-step finds the parameter estimates to maximize the complete-data log likelihood from the E-step. The two steps are iterated until the iterations converge.

[8] After re-normalization of the pillar scores, zero values were replaced with 0.00001 to avoid that zero values in one pillar result in CRI scores equal to 0.0 regardless of the country's performance on the other two pillars.

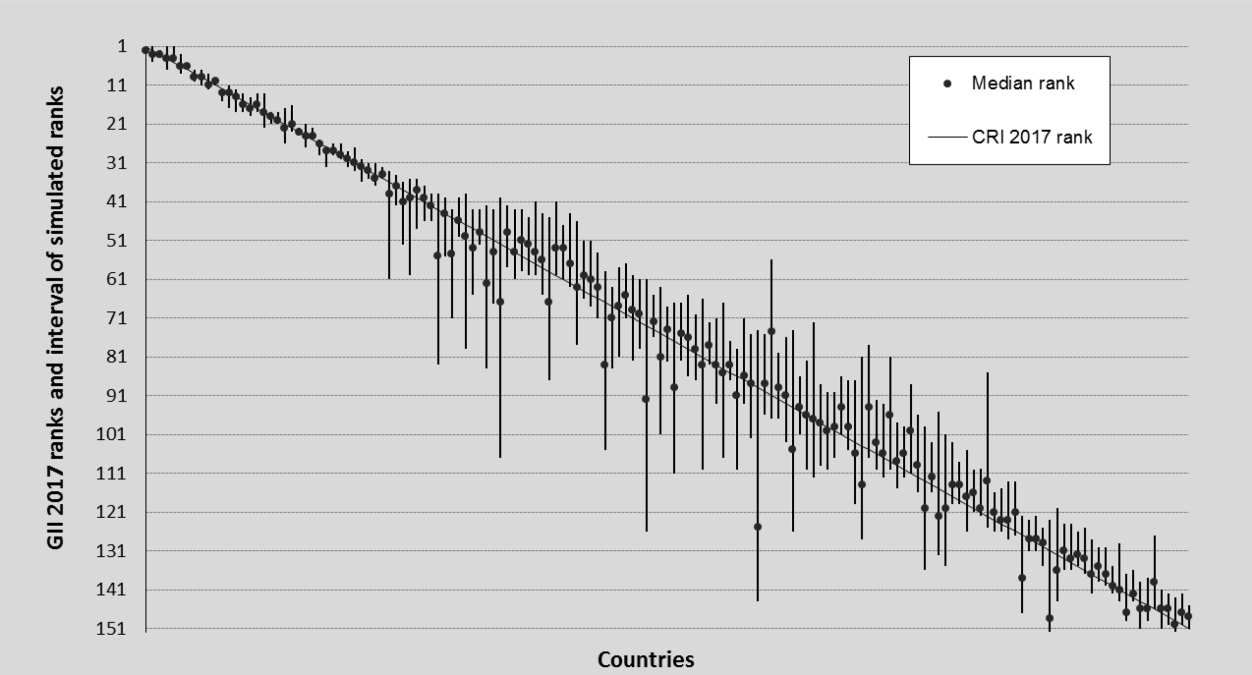
Four models were tested based on the combination of no imputation *versus* EM imputation, and original aggregation formula *versus* geometric average of the three pillars. A total of 4,000 simulations were carried out combining the four models with the 1,000 simulations *per* model corresponding the 1000 different sets of weight assigned to each of the three pillars. Table 6 summarizes the uncertainties considered for the robustness assessment of the CRI index.

Uncertainty analysis results

The robustness analysis results for the 152 countries are summarised in Figure 2 with median ranks and 90% confidence intervals computed across the 4,000 Monte Carlo simulations of the CRI overall score. Countries are ordered from best to worst according to their reference rank (black line), the dot being the median rank. Error bars represent, for each country, the 90% interval across all simulations. Table 7 reports the published rankings and the 90% confidence intervals that account for uncertainties in the missing data estimation, the pillar weights, and the aggregation formula. All published country ranks lay within the simulated intervals, and the fact that the CRI rank is close to the median rank for the majority of the countries suggests that the CRI is a suitable summary measure.

CRI ranks are shown to be both representative of a plurality of scenarios and robust enough (for about one-third of the countries) to changes in the imputation method, the pillar weights, and the aggregation formula. If one considers the median rank across the simulated scenarios as being representative of these scenarios, then the fact that the CRI rank is close to the median rank (less than 4 positions away) for 75% of the countries suggests that the CRI is a suitable summary measure. Furthermore, the narrow confidence intervals for one-third of the countries' ranks (less than ± 5 positions) imply that for those countries – mostly OECD countries – the CRI ranks are robust enough to changes in the pillar weights, the imputation method, and the aggregation formula. Nevertheless, caution is needed for a number of countries whose CRI rank is sensitive to the computation methodology. For full transparency and information, Table 7 reports the CRI country ranks together with the simulated intervals (90% of the 4,000 scenarios) and the median rank across all simulations in order to better appreciate the robustness of these ranks to the computation methodology.

Figure 2: Robustness analysis (CRI rank vs. median rank, 90% confidence intervals)



Source: European Commission, Joint Research Centre (JRC), 2017.
Notes: The Spearman rank correlation between the median rank and the CRI rank is 0.99. Median ranks and intervals are calculated over 4,000 simulated scenarios combining random weights at the pillar level, imputed versus missing values, and geometric versus original aggregation formula at the pillar level.

Table 7: Country ranks and 90% intervals for the CRI 2017

		Simulations				Simulations				Simulations	
	Rank	Median	Interval		Rank	Median	Interval		Rank	Median	Interval
Sweden	1	2	[1, 2]	Romania	51	54	[43, 67]	Barbados	101	94	[86, 101]
Norway	2	3	[1, 5]	Guinea	52	67	[40, 107]	Rwanda	102	99	[87, 105]
Denmark	3	3	[2, 4]	Bolivia	53	49	[42, 58]	Morocco	103	106	[87, 119]
Belgium	4	4	[1, 7]	Djibouti	54	54	[43, 61]	Cambodia	104	114	[81, 128]
Germany	5	4	[1, 5]	Zimbabwe	55	51	[43, 59]	Antigua and Barbuda	105	94	[78, 107]
Austria	6	6	[3, 8]	Mauritius	56	52	[45, 60]	Ethiopia	106	103	[92, 110]
Finland	7	6	[6, 7]	Malawi	57	54	[41, 60]	Ghana	107	106	[93, 112]
Netherlands	8	9	[7, 10]	Korea, Rep.	58	56	[44, 65]	St. Vincent and the Grenadines	108	96	[81, 110]
France	9	9	[7, 11]	Kiribati	59	67	[45, 87]	Guatemala	109	108	[98, 115]
Japan	10	11	[8, 12]	Seychelles	60	53	[41, 60]	Congo, Dem. Rep.	110	106	[99, 112]
Luxembourg	11	10	[10, 11]	Turkey	61	53	[47, 61]	Tonga	111	100	[88, 111]
Iceland	12	13	[12, 15]	Chile	62	57	[44, 63]	Congo, Rep.	112	109	[96, 116]
Australia	13	13	[11, 17]	Togo	63	63	[46, 78]	Malaysia	113	120	[99, 136]
Ireland	14	14	[12, 18]	Armenia	64	60	[51, 66]	Indonesia	114	112	[103, 116]
Canada	15	16	[13, 18]	Jordan	65	61	[51, 68]	Central African Republic	115	122	[95, 132]
United Kingdom	16	17	[14, 19]	Tajikistan	66	63	[54, 71]	Botswana	116	120	[101, 135]
Italy	17	16	[13, 18]	Solomon Islands	67	83	[59, 105]	Cote d'Ivoire	117	114	[103, 119]
Switzerland	18	18	[13, 22]	Maldives	68	71	[63, 84]	Gambia, The	118	114	[108, 119]

Slovenia	19	19	[18, 21]	Sierra Leone	69	68	[58, 81]	Azerbaijan	119	117	[105, 126]
Portugal	20	20	[18, 21]	Benin	70	65	[57, 71]	Philippines	120	116	[110, 121]
South Africa	21	22	[17, 26]	El Salvador	71	69	[60, 82]	Uganda	121	120	[110, 122]
Malta	22	21	[16, 23]	Dominican Republic	72	70	[61, 79]	Moldova	122	113	[85, 125]
United States	23	23	[22, 24]	Thailand	73	92	[61, 126]	Senegal	123	121	[116, 126]
Czech Republic	24	24	[21, 27]	Trinidad and Tobago	74	72	[65, 76]	Jamaica	124	123	[115, 126]
Argentina	25	24	[22, 25]	Peru	75	81	[63, 101]	Angola	125	123	[113, 128]
Greece	26	26	[25, 29]	Mongolia	76	74	[68, 82]	Bhutan	126	121	[113, 127]
Spain	27	28	[27, 32]	Sao Tome and Principe	77	89	[67, 111]	Panama	127	138	[122, 147]
Hungary	28	28	[26, 29]	Nepal	78	75	[67, 82]	Fiji	128	128	[123, 131]
Israel	29	29	[26, 30]	Paraguay	79	76	[65, 86]	Guinea-Bissau	129	128	[122, 131]
New Zealand	30	30	[28, 32]	Vietnam	80	79	[70, 87]	Yemen, Rep.	130	129	[125, 135]
Cyprus	31	31	[27, 33]	Bulgaria	81	83	[66, 110]	Egypt, Arab Rep.	131	149	[123, 152]
Slovak Republic	32	32	[30, 36]	St. Lucia	82	78	[72, 83]	Madagascar	132	136	[120, 144]
Croatia	33	33	[31, 35]	Kazakhstan	83	83	[71, 93]	Cabo Verde	133	131	[124, 136]
Costa Rica	34	35	[31, 37]	Lithuania	84	85	[67, 107]	India	134	133	[124, 136]
Poland	35	34	[32, 35]	Algeria	85	83	[77, 87]	Lebanon	135	132	[126, 135]
Liberia	36	39	[33, 61]	Occupied Palestinian Territories	86	91	[79, 110]	Cameroon	136	133	[125, 137]
Uruguay	37	37	[34, 42]	Tanzania	87	86	[71, 93]	Oman	137	137	[128, 142]
Samoa	38	41	[36, 52]	Zambia	88	88	[75, 102]	Haiti	138	135	[130, 139]
Niger	39	40	[35, 60]	Georgia	89	125	[74, 144]	Pakistan	139	137	[130, 140]
Estonia	40	38	[35, 48]	Mexico	90	88	[75, 96]	Sri Lanka	140	140	[135, 142]
Namibia	41	40	[37, 46]	Serbia	91	75	[56, 97]	Ukraine	141	141	[129, 144]
Guyana	42	42	[39, 46]	China	92	89	[80, 97]	Swaziland	142	147	[137, 149]
Tunisia	43	55	[39, 83]	Honduras	93	91	[76, 103]	Bangladesh	143	142	[136, 144]
Colombia	44	44	[40, 55]	Russian Federation	94	105	[74, 126]	Vanuatu	144	146	[139, 151]
Lesotho	45	55	[43, 71]	Singapore	95	94	[86, 101]	Myanmar	145	146	[138, 149]
Mozambique	46	46	[40, 50]	Kenya	96	96	[82, 110]	Lao PDR	146	139	[127, 146]
Latvia	47	50	[39, 79]	Mauritania	97	97	[72, 112]	Belarus	147	146	[141, 151]
Burkina Faso	48	53	[43, 65]	Kyrgyz Republic	98	98	[88, 109]	Afghanistan	148	146	[142, 150]
Ecuador	49	49	[43, 52]	Burundi	99	100	[90, 110]	Albania	149	150	[143, 152]
Papua New Guinea	50	62	[42, 84]	Mali	100	99	[90, 107]	Timor-Leste	150	147	[142, 150]

Source: European Commission, Joint Research Centre (JRC), 2017.Notes: Median ranks and intervals are calculated over 4,000 simulated scenarios combining random weights, imputed versus missing values, and geometric versus original aggregation formula at the pillar level.

Sensitivity analysis results

To complement the uncertainty analysis, sensitivity analysis has been used in order to identify which of the modelling assumptions tested previously has the highest impact on country ranks. Table 8 summarizes the impact of estimating the missing data with the EM imputation method as well as the effect of adopting a geometric aggregation formula (assuming equal weights for the three pillars as in the published CRI).

When the geometric average is used across all three pillars, five countries – Georgia, Guinea, Russian Federation, Thailand, Tunisia – decline by more than 20 positions, while no country improves by 10 or more positions. The impact of estimating missing data is more noteworthy: two countries – Antigua and Barbuda, Serbia – improve by 20 positions or more, while six countries – Georgia, Lesotho, Occupied Palestinian Territory, Papua New Guinea, São Tomé and Principe, Solomon Islands – decline by 20 positions or more. The combination of these two assumptions, namely the EM estimation for missing data and the geometric average of the three pillars has a more pronounced effect. Yet, these assumptions concern methodological choices only and might overall be less influential than choices related to the background assumptions in the conceptual framework (Saltelli and Funtowicz, 2014).

All in all, the published CRI ranks are reliable and representative of a plurality of methodological scenarios for most countries. Furthermore, for one-third of the countries the simulated 90% rank intervals are narrow enough for meaningful inferences to be drawn. The readers of the CRI report should consider country ranks not only at face value but also within the 90% confidence intervals in order to better appreciate to what degree a country's rank depends on the modelling choices.

Table 8: Sensitivity analysis: Impact of modelling choices on countries with most sensitive ranks

	Number of countries that improve				Number of countries that deteriorate			
Uncertainty tested	**by 20 or more positions**		**between 10-19 positions**		**by 20 or more positions**		**between 10-19 positions**	
Geometric average versus original CRI aggregation formula (pillar level)	0		0		5	Georgia, Guinea, Russian Federation, Thailand, Tunisia	9	Arab Rep., Botswana, Burkina Faso, Cambodia, Central African Republic, Egypt, Malaysia, Panama, Peru
EM imputation vs. no imputation of missing data (12 indicators dataset)	2	Antigua and Barbuda, Serbia	3	Moldova, St. Vincent and the Grenadines, Tonga	6	Georgia, Lesotho, Occupied Palestinian Territories, Papua New Guinea, Sao Tome and Principe, Solomon Islands	3	Kiribati Maldives Samoa

Geometric average and EM imputation vs. original CRI aggregation formula and no estimation of missing values	5	Antigua and Barbuda, Moldova, Serbia, St. Vincent and the Grenadines, Tonga	7	Barbados, Benin, Bhutan, Chile ,Lao PDR, Seychelles, Turkey	8	Georgia, Guinea, Lesotho, Papua New Guinea, Russian Federation, Sao Tome and Principe, Solomon Islands, Thailand	12	Botswana, Cambodia, Central African Republic, Egypt, Arab Rep., Kiribati, Liberia, Malaysia, Occupied Palestinian Territories, Panama, Peru, Tunisia

Source: European Commission Joint Research Centre, 2017.

3. Conclusion

The JRC analysis suggests that the conceptualized multi-level structure of the CRI index is statistically coherent and balanced (i.e. not dominated by any pillar, and all eight indicators contributing to the variation of their respective pillar scores). The reliability of the CRI, as measured by the Cronbach-alpha value is good at 0.71 (just above the recommended 0.7 threshold for a reliable aggregate).

Points that call for possible refinements of the CRI framework were also identified. These refinements regard mainly two out of the eight indicators, namely 'Tax structure' under the Tax pillar and 'Minimum Wage as a % GDP' under the Labour pillar, which were found to be not influential at the overall CRI level though they are at their own pillar level. This is indicative of a different behaviour of these two indicators with respect to the remaining ones composing the CRI index. The JRC recommends to the CRI developing team to keep these indicators in the current framework because of their conceptual relevance to the phenomenon but to test and eventually refine next year's release along these issues if next year's data confirm the same pattern.

The CRI country ranks are robust to methodological assumptions related to the estimation of missing data, weighting and aggregation formula. It is reassuring that for 75% of the 152 countries, the CRI rank is close (less than 4 positions away) to the median rank calculated over 4,000 simulations (combinations of modelling choices related to the estimation of missing data, the pillar weights and the aggregation formula at the pillar level). Furthermore, for one-third of the countries, the confidence intervals are narrow enough to allow for inferences to be drawn. Caution however is needed for non-OECD countries whose rank is more sensitive to the methodological choices. Note that a high robustness in the case of the CRI would have been undesirable as this would have implied that the three pillars are perfectly correlated and hence redundant.

One way in which the CRI helps to highlight what governments in 152 countries are doing to tackle the growing gap between rich and poor is by pinpointing the differences in rankings that emerge from a comparison between the CRI overall index and each of the three pillars: for more than 60% (up to 74%) of the 152 countries included in the CRI Index, the CRI ranking and any of the three pillar rankings differ by 10 positions or more. This outcome evidences both the added value of the CRI ranking and points at the importance of duly taking into account the individual pillars and indicators on their own merit. By doing so, country-specific strengths and bottlenecks on reducing inequality can be identified and serve as an input for evidence-informed policymaking.

The auditing conducted herein has shown the potential of the CRI in paving the way towards a monitoring framework that can help to identify weaknesses and best practices in governments' efforts to reduce the gap between the rich and the poor and ultimately guide policy formulation and action.

4. References and related reading

Becker, W., Saisana, M., Paruolo, P., Vandecasteele, I. (2017) Weights and importance in composite indicators: Closing the gap, Ecological Indicators 80: 12–22.

Groeneveld, R.A. and Meeden, G. 1984. Measuring skewness and kurtosis. The Statistician 33, 391–399.

Little, R.J.A., Rubin, D.B. 2002. Statistical Analysis with missing data. 2nd edition; John Wiley & Sons, Inc.

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

Nunnaly, J. 1978. Psychometric theory, New York: McGraw-Hill.

OECD/EC JRC. 2008. Handbook on Constructing Composite Indicators: Methodology and User Guide. Paris: OECD.

Paruolo, P., Saisana, M., Saltelli, A. 2013. Ratings and rankings: Voodoo or Science?, Journal of the Royal Statistical Society – A 176(3), 603–634.

Saisana, M., D'Hombres, B., Saltelli, A. 2011. Rickety numbers: Volatility of university rankings and policy implications, Research Policy 40, 165–177.

Saisana, M., Saltelli, A. 2011. Rankings and Ratings: Instructions for use, Hague Journal on the Rule of Law 3(2), 247–268.

Saisana, M., Saltelli, A., Tarantola, S., 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), 307–323.

Saltelli, A., Funtowicz, S. 2014. When all models are wrong. Issues in Science and Technology, Winter 2014, 79–85.

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

Published by Oxfam for JRC in July 2017. Oxfam GB, Oxfam House, John Smith Drive, Cowley, Oxford OX4 2JY, UK.