European Commission

JRC TECHNICAL REPORTS

JRC statistical audit of Commitment to Reducing Inequality index 2018

Álvarez-Díaz, Marcos
Saisana, Michaela
D'Hombres, Beatrice
Montalto, Valentina
Tacao Moura, Carlos

2018

Joint Research Centre

EUR 29406 EN

Contact information
Competence Centre on Composite Indicators and Scoreboards
jrc-coin@ec.europa.eu
https://ec.europa.eu/jrc/en/coin
https://composite-indicators.jrc.ec.europa.eu/

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

JRC113227

EUR 29406 EN

PDF	ISBN 978-92-79-96744-3	ISSN 1831-9424	doi:10.2760/942703
Print	ISBN 978-92-79-96743-6	ISSN 1018-5593	doi:10.2760/83485

Luxembourg: Publications Office of the European Union, 2018

How to cite this report: Alvarez Diaz, M., Saisana, M., D`hombres, B., Montalto, V. and Tacao Moura, C., *JRC Statistical Audit of Commitment to Reducing Inequality Index 2018*, EUR 29406 EN, Publications Office of the European Union, Luxembourg, 2018, ISBN 978-92-79-96744-3, doi:10.2760/942703, JRC113227.

Contents

Abstract

Eradicating poverty is one of the global challenges defined in the 2030 Agenda for Sustainable Development. However, this goal cannot be achieved without reducing the gap between the rich and the poor. Development Finance International and Oxfam have developed an international monitoring framework — the Commitment to Reducing Inequality (CRI) index — that measures the commitment of 157 countries to reducing inequality through the fiscal policies (public spending and taxes) and labour market policies implemented by their governments. The CRI index builds on three pillars: progressivity of spending, progressivity of tax, and progressivity of labour policy. These pillars are used to organise and aggregate nine indicators into a single summary measure. This framework involves both conceptual and practical challenges. The statistical audit presented here was performed by the European Commission's Joint Research Centre, and it aims to contribute to ensuring the transparency and reliability of the CRI index 2018. It should enable policymakers to derive more accurate and meaningful conclusions, and to potentially guide choices on priority setting and policy formulation.

Overall, the main conclusions of the present audit can be summarised as follows: the CRI index 2018 is representative of a plurality of scenarios, is reliable and has a statistically coherent framework. The uncertainty analysis shows that country ranks are robust for most countries. For a number of countries, in particular countries that are not members of the Organisation for Economic Cooperation and Development (OECD), ranks should be analysed within their expected confidence intervals instead of being taken at face value. The statistical assessment also shows that the CRI index has a good statistical reliability and measures one single latent phenomenon capturing the main components of the index: the 'progressivity of labour policies', and the interaction between the 'progressivity of tax' and the 'progressivity of spending'. Notwithstanding the good statistical properties of the CRI index, some suggestions are made for possible refinements of the CRI index in future editions.

1. Introduction

The Commitment to Reducing Inequality (CRI) index aims at measuring the extent to which governments are undertaking the task of reducing inequality taking into account three main aspects: progressivity of spending, progressivity of tax and progressivity of labour policy. These pillars are used to organise and aggregate nine indicators into a single summary measure. The index is developed by Development Finance International and Oxfam

The main report discusses in detail the conceptual framework of the CRI index 2018, as well as the selection of indicators, data quality aspects and methodological choices for grouping country-level data. The CRI index 2018 is composed of nine indicators and three pillars. The overall score of the index is calculated as the weighted sum of two thirds of the square root of the interaction between the progressivity of spending and progressivity of tax pillars, plus one third of the score for the labour pillar. Table 1 summarises the conceptual framework of the CRI index 2018, including the aggregation rule. The CRI framework is well constructed and a lot of thought has clearly been put into it. However, conceptual and practical challenges are inevitable when trying to summarise with a single composite indicator the commitment of countries to reducing inequality. An analysis is needed to ensure and validate the statistical soundness of any composite index.

This audit was performed by the European Commission's Competence Centre on Composite Indicators and Scoreboards at the Joint Research Centre (JRC), and was conducted upon the invitation of the developers. The analysis herein aims at shedding light on the transparency and reliability of the CRI index 2018 and thus to enabling policymakers to derive more accurate and meaningful conclusions, and to potentially guide choices on priority setting and policy formulation.

Table 1: *CRI index 2018: conceptual framework, weights and aggregation*

OVERALL INDEX

$$\boldsymbol{CRI\ 2018\ index} = \frac{\mathbf{2} \cdot \sqrt{\boldsymbol{progressivity\ of\ spending} \cdot \boldsymbol{progressivity\ of\ tax}} + \boldsymbol{progressivity\ of\ labour}}{\mathbf{3}}$$

INDICATORS

Progressivity of spending [S]		Progressivity of tax [T]		Progressivity of labour policy [L]	
Su-pillar	**Weight**	**Indicator**	**Weight**	**Indicator**	**Weight**
S1 social spending as % of total spending	Weight 50 %	**T1 Progressivity of tax structure**	25 %	**L1 Workers and labour union rights**	33 %
S2 incidence of spending on inequality (Gini coefficient)	Weight 50 %	**T2 Incidence of tax on inequality (Gini coefficient)**	25 %	**L2 Women's legal rights at work**	33 %
		T3 Tax collection	25 %		
		T4 Harmful Tax Practices	25 %	**L.3 Minimum wage**	33 %

Source: Commitment to Reducing Inequality index 2018.

In general, 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 discussions. Suggestions for fine tuning, aimed at setting the foundation for a balanced index, were taken into account by Oxfam and the Development Finance International research teams for the final computation of the CRI scores and rankings.

The JRC assessment of the CRI index presented here focuses on two main issues: the statistical coherence of the structure, and the impact of key modelling assumptions on the CRI scores and ranks ([2]). The statistical analysis is based on the adequacy of aggregating indicators into pillars, and pillars into the overall index. Nevertheless, given the particular aggregation formula used by the developing team, the analysis is also focused on the aggregation of the labour pillar and the interaction between the spending and tax pillars. Finally, the JRC analysis complements the reported country rankings for the CRI index 2018 with estimated confidence intervals, in order to better appreciate the robustness of these ranks to some modelling choices (such as the weighting scheme, aggregation formula and estimation of missing values).

([1]) OECD and JRC (2008).

([2]) The JRC analysis was based on the recommendations of the OECD and JRC Handbook on Composite Indicators (2008), 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 and at https://composite-indicators.jrc.ec.europa.eu/ (all audits were conducted upon request of the Index developers).

2. Statistical coherence in the Commitment to Reducing Inequality index framework

The variables used to derive the nine indicators and three pillars that compose the overall CRI index were selected by Oxfam and the Development Finance International research teams for their relevance to 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 annex. The assessment of the statistical coherence of the CRI index starts from the level of the nine main indicators grouped across the three pillars and further aggregated into the overall CRI index. The overall index is calculated as the weighted sum of the geometric average of the spending and tax pillars, and the score of the labour pillar.

The present statistical assessment of the CRI index involves the following three steps. In the first step, the main descriptive statistics of the data are shown, and an initial analysis of the data is performed to detect missing values and potential outliers. In the next step, the statistical coherence is examined through a multilevel analysis of the correlations of the indicators and pillars. Moreover, the importance of each indicator is explored. Finally, in the last step, the added value of the CRI index 2018 is studied through the comparison of the CRI ranking with the ranking of its pillars, and with other indices measuring the current level of country inequality.

2.1. Data checks

Data coverage has increased with respect to the previous edition. The CRI index 2018 has been calculated for 157 countries. This coverage implies five additional countries in comparison with the last edition (Brazil, Belize, Chad, Kosovo ([3]) and Uzbekistan). Additionally, the Development Finance International team identified and included in the current edition of the index more reliable and recent data sources. This is an important point given that the quality and adequacy of the index lies not only on the index development, but also on getting reliable data.

Table 2 offers summary statistics for the indicators and pillars. The three indicators composing the progressivity of spending pillar do not include any missing values in the final data set. For the second pillar, progressivity of tax, the tax collection effort indicator presents only one missing value (occupied Palestinian territories), and the harmful tax practices indicator has 27 missing values (17.2 % of the countries). The third pillar, progressivity of labour policies, only has three missing values for the workers and labour union rights indicator (Bhutan, Kosovo and Tonga). This implies that the overall CRI scores could be biased for those countries with missing values. In the uncertainty and sensitivity analysis, the robustness of the CRI ranking to an alternative imputation method will be tested.

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 treated if the indicators have absolute skewness greater than 2.0 and kurtosis greater than 3.5 ([4]). A potential problem arises in the minimum wage indicator. As shown in Table 2, the skewness and kurtosis for this indicator are extremely high (5.75 and 53.35, respectively). The Central African Republic presents an extremely high value for minimum wage (see Figure 1). The existence of this outlier can cause possible misinterpretations of

([3]) This designation is without prejudice to positions on status, and is in line with UNSCR 1244/1999 and the ICJ Opinion on the Kosovo declaration of independence.

([4]) Groeneveld and Meeden (1984) set the criteria for absolute skewness above 1 and kurtosis above 3.5.

the scores and rankings of the index. For that reason, it is recommended for future editions of the CRI index to detect and correct outliers.

Table 2 also reflects that the social spending and tax structure indicators were not normalised following the min/max formula. This decision could underestimate the effect of these indicators on their respective pillars and, consequently, on the overall index. Thus, an additional recommendation for future editions is to normalise these indicators to cover the whole range [0, 1] ([5]).

Table 2: *Summary statistics of indicators and pillars*

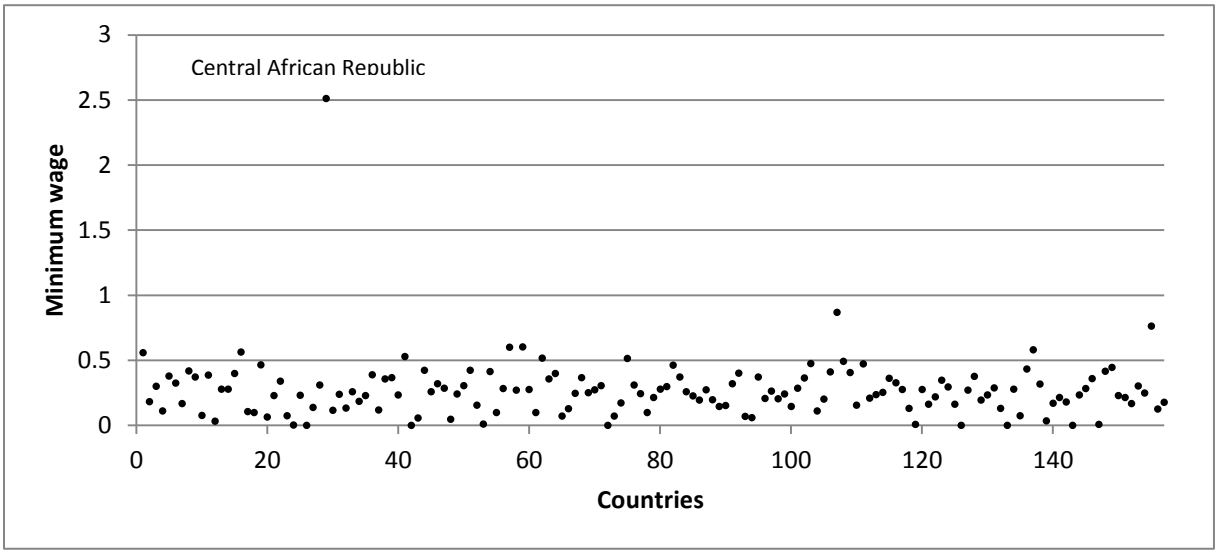
CRI components	Number of observations	Missing data (%)	Mean	skewness	kurtosis	[Min, max]
Pillar 1: progressivity of spending [S]	**157**	**0.00 %**	**0.32**	**0.87**	**−0.23**	**[0, 1]**
Social spending [S.1]	157	0.00 %	0.35	0.17	−0.87	[0.06, 0.72]
Incidence of spending on inequality [S.2]	157	0.00 %	0.16	1.63	2.29	[0, 1]
Pillar 2: progressivity of tax [T]	**157**	**0.00 %**	**0.52**	**−0.23**	**0.10**	**[0, 1]**
Tax structure [T.1]	157	0.00 %	0.58	−0.58	0.30	[0.12, 0.93]
Tax incidence [T.2]	157	0.00 %	0.62	−0.81	2.58	[0, 1]
Tax collection [T.3]	156	0.64 %	0.44	0.31	−0.26	[0, 1]
Harmful tax practices [T.4]	130	17.20 %	0.72	−1.42	3.48	[0, 1]
Pillar 3: progressivity of labour policies [L]	**157**	**0**	**0.46**	**0.23**	**−0.99**	**[0, 1]**
Workers and labour union rights [L.1]	154	1.91 %	0.43	0.28	−1.02	[0, 1]
Women's legal rights at work [L.2]	157	0.00 %	0.47	0.12	−1.18	[0, 1]
Minimum wage [L.3]	157	0.00 %	0.11	5.59	50.82	[0, 1]

Source: European Commission, Joint Research Centre (JRC), 2018.
NB: Rule for outlier detection: indicators with a |skewness| >2 and a kurtosis >3.5.

The missing value for the tax collection indicator is occupied Palestinian territories; the missing values for the harmful tax practices indicator are Bangladesh, Benin, Bhutan, Burkina Faso, Burundi, Central African Republic, Chad, Djibouti, Ethiopia, Guinea, Guinea-Bissau, Haiti, Kiribati, Lesotho, Mali, Mauritania, Nepal, Niger, occupied Palestinian territories, São Tomé and Príncipe, Sierra Leone, Solomon Islands, The Gambia, Togo, Tonga, Yemen and Zambia; finally, the outliers for the workers and labour union rights indicator are Bhutan, Kosovo and Tonga.

([5]) The JRC team also performed the statistical analysis of the CRI index 2018 treating the outlier through Winsorisation and normalising the social spending and tax structure indicators. The statistical coherence of the index was not significantly modified, but some changes in the score and ranking of the countries were observed.

Figure 1: *Central African Republic's outlier performance in the minimum wage indicator*



Source: European Commission, Joint Research Centre (JRC), 2018.

2.2. Statistical coherence

The reliability of the CRI index 2018 depends — among other things — on the degree of coherence between the conceptual framework and the statistical structure of the data. The more the conceptual framework embraces the statistical structure, the higher the reliability of the CRI index will be. The coherence of the framework was assessed using two tests: (a) analysing the extent to which the indicators can explain a sufficient amount of variation in the aggregated scores (be those indicators, pillars or the overall index) by means of cross-correlation; and (b) analysing the importance of the indicators in the CRI framework.

Cross-correlation analysis

The cross-correlation analysis was used to assess to what extent the data support the conceptual framework. Table 4 shows the correlation between the different components of the CRI index. A 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 with any of the other two pillars. This suggests that the allocation of the indicators to the specific pillar is consistent both from conceptual and statistical perspectives. Nevertheless, there is some suspicion of redundant information in the progressivity of labour policies pillar. This suspicion is based on the high correlation that the workers and labour union rights and women's rights in the workplace indicators have with the pillar to which they belong (Pearson correlation coefficients greater than 0.92).

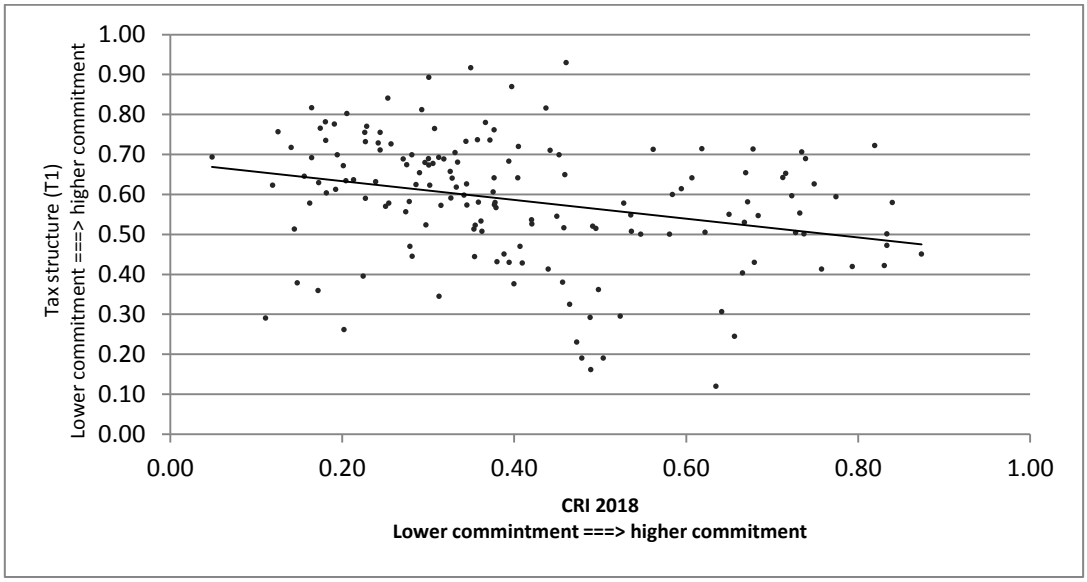
Table 3: Pairwise correlations between indicators, pillars and the CRI index

	Pillar 1: progressivity of spending [S]	**Pillar 2: progressivity of tax [T]**	**Pillar 3: progressivity of labour policies [L]**	**CRI index**
Social spending [S.1]	0.90	0.26	0.72	0.83
Incidence of spending on inequality [S.2]	0.94	0.36	0.71	0.85
Tax structure [T.1]	−0.35	0.23	−0.42	−0.29
Tax incidence [T.2]	0.27	0.49	n.s	0.32
Tax collection [T.3]	0.51	0.55	0.49	0.63
Harmful tax practices [T.4]	n.s	0.37	n.s	n.s
Workers and labour union rights [L.1]	0.72	n.s	0.93	0.82
Women's rights in the workplace [L.2]	0.78	0.30	0.92	0.87
Minimum wage [L.3]	n.s	n.s	0.37	0.23

Source: European Commission, Joint Research Centre (JRC), 2018.
NB: Numbers represent Pearson correlation coefficients. 'n.s': non-significant correlations at the 1 % level.

The correlations between the indicators and the CRI index reveal that four out of the nine indicators — social spending, incidence of spending, workers and labour union rights and women's legal rights at work — show a strong, positive and significant correlation with the overall CRI index (correlation greater than 0.82). Three indicators — tax collection, tax incidence and minimum wage — are moderately, but significantly, correlated with the CRI index. The harmful tax practices indicator does not seem to be influential at the overall CRI index level, though it is moderately correlated with its own pillar level. More problematic is the unexpected negative and significant correlation observed between the tax structure indicator and the CRI index. This association means that those countries with higher and more progressive direct tax rates and lower indirect taxes experience a lower commitment to reducing inequality. The negative association is depicted in Figure 1. This statistical finding seems to go against the theoretical framework behind the CRI index and is worthy of further reflection by the CRI index developers.

Figure 2: *Relationship between the tax structure indicator and the CRI index 2018*



Source: European Commission, Joint Research Centre (JRC), 2018.

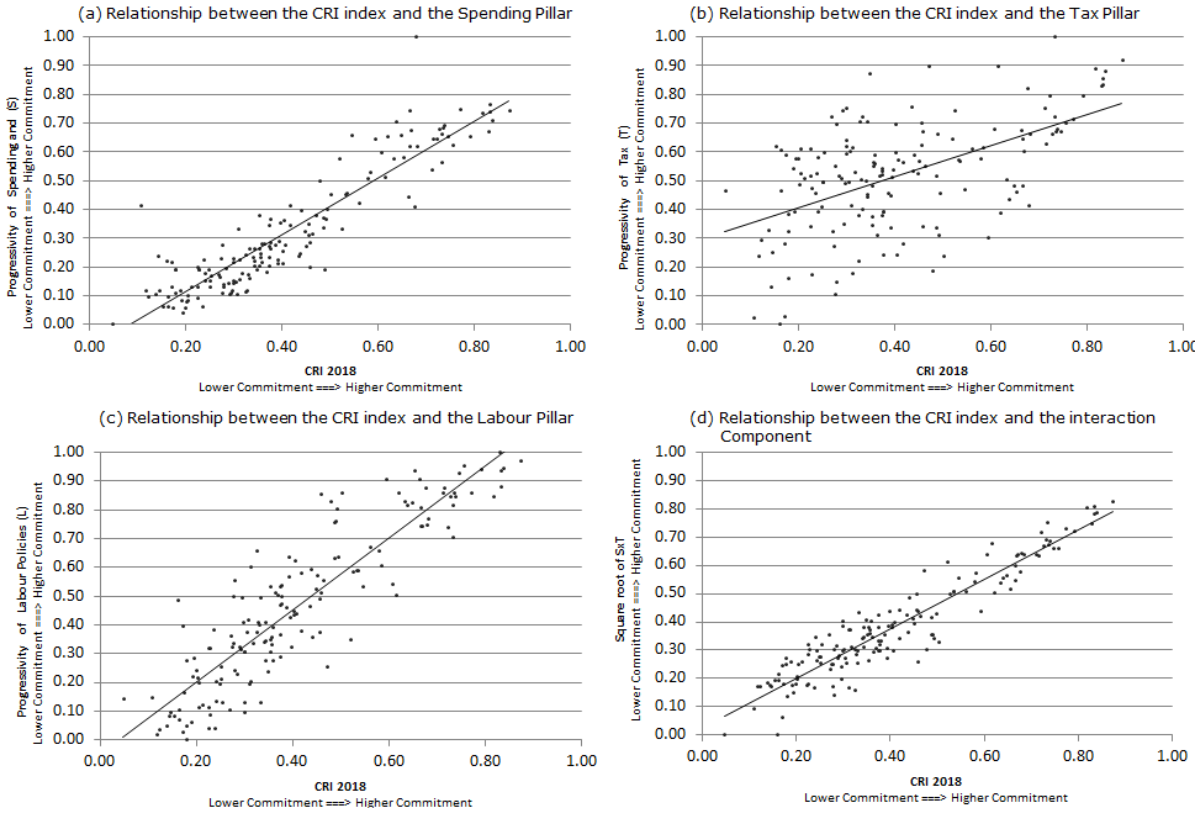
At the overall CRI index level, the three pillars correlate strongly with the CRI index, with pairwise correlations above 0.57 (see Table 4). The spending and labour pillars are, however, more strongly associated with the overall CRI index than the tax pillar. The component that captures the interaction between the spending and tax pillars ($(\sqrt{\boldsymbol{spending} \cdot \boldsymbol{tax}})$) strongly correlates with the index (coefficient of correlation equal to 0.94). Furthermore, the aggregation of the spending and tax pillars into one component seems to be adequate from a statistical point of view. This comes from the fact that the correlation coefficients of these pillars are high (greater than 0.7) but not so excessive as to cause a problem of redundant information (less than 0.92). Figure 3 describes the strong linear association between the pillars and the CRI index (see Figures 1(a) to 1(c)), as well as the association with the interaction component (see Figure 1(d)).

Table 4: *Pairwise correlations between pillars and the CRI index*

	Spending pillar (S)	**Tax pillar (T)**	**Labour pillar (L)**	**CRI index**
Spending pillar (S)	1.00	0.35	0.77	0.91
Tax pillar (T)	0.35	1	n.s	0.57
Interaction component $(\sqrt{\boldsymbol{spending} \cdot \boldsymbol{tax}})$	0.89	0.72	0.64	0.94
Labour pillar (L)	0.77	0.23	1	0.89

Source: European Commission, Joint Research Centre (JRC), 2018.
NB: Numbers represent Pearson correlation coefficients.

Figure 3: *Relationship between the pillars, the interaction component and the CRI index*



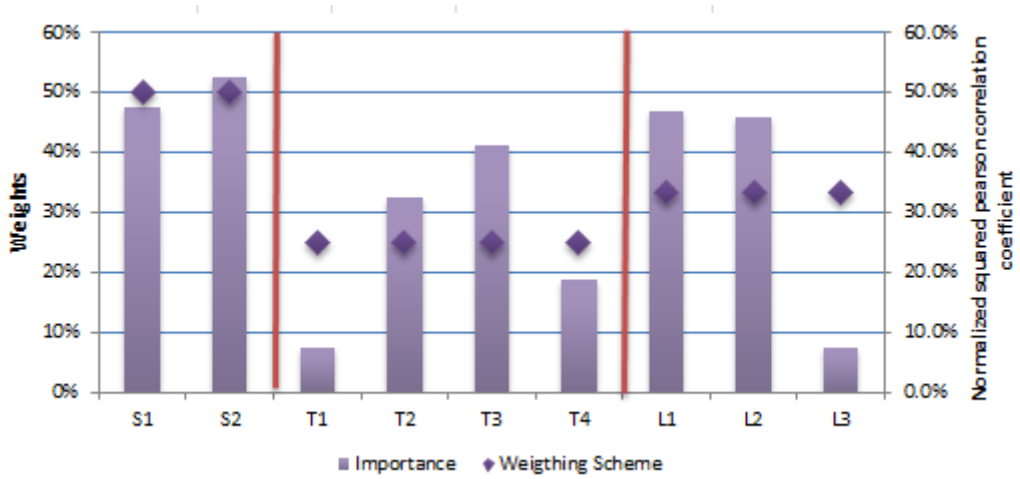
Source: European Commission, Joint Research Centre (JRC), 2018.

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 nine indicators at various levels of aggregation in the CRI structure. We use the normalised squared Pearson correlation coefficients as a statistical measure of the importance of indicators. 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 4. The dots correspond to the weights assigned to each indicator within the pillar to which the indicator belongs whilst the bars represent the actual statistical importance of the indicators.

Figure 4: *Weights and statistical importance of the indicators within each pillar*

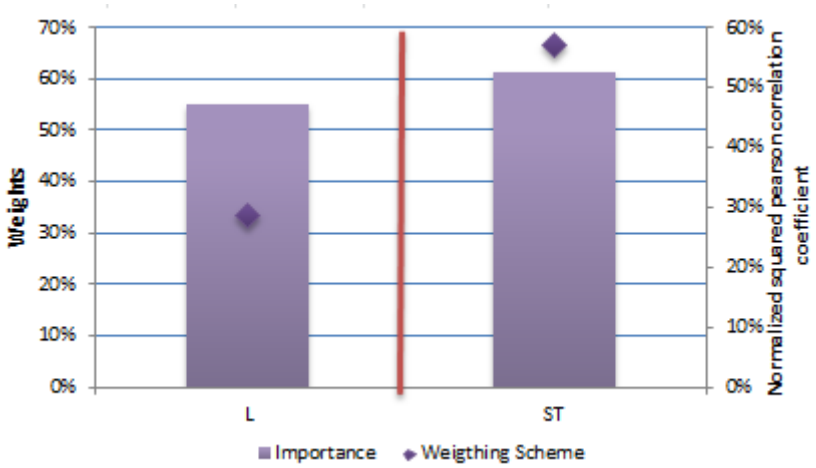


Source: European Commission, Joint Research Centre (JRC), 2018.

Within the first pillar, the indicator on social spending (S1) captures a slightly smaller proportion of the variance of the spending pillar compared to the indicator measuring the incidence of spending on inequality (S2) (47.6 % versus 52.4 %). Given that an equal weight was applied to the two indicators, this implies that the weighting scheme adequately captures their actual importance. For the second pillar on the progressivity of tax, the tax incidence (T2) and tax collection (T3) indicators are important to explain the variance of the tax pillar (32.6 % and 41.2 %, respectively). Conversely, the harmful tax practices (T4) and tax structure (T1) indicators have little importance (18.9 % and 7.3 %, respectively). Regarding the pillar on the progressivity of labour policies, the workers and labour union rights (L1) and women's legal rights at work (L2) indicators are much more influential (46.8 % and 45.8 %, respectively) than the third indicator measuring the fairness of the level of minimum wages (L3) (7.4 %). The analysis of the importance of the indicators in the CRI framework is in line with the correlation analysis results, and suggests that the CRI developing team might need to reconsider how best to include the harmful tax practices (T4), tax structure (T1) and minimum wages (L3) indicators in next year's release, should this statistical result also be confirmed with next year's data set. For this release, it is only being flagged for attention.

The CRI index 2018 is calculated based on the weighted average of two components: the geometric aggregation of the spending and tax pillars, and the score of the labour pillar. As already mentioned, the weights assigned for each component are two thirds for the interaction component and one third for the labour pillar component. Figure 5 shows that the labour pillar component and the interaction component are placed on similar footing (47.3 and 52.7 %, respectively).

Figure 5: *Weights and statistical importance of the interaction component and labour pillar within the CRI index*



Source: European Commission, Joint Research Centre (JRC), 2018.

2.3. Added value of the Commitment to Reducing Inequality index

For 61 % or more of the 157 countries included in the CRI index, the CRI ranking and any of the three pillar rankings differ by 10 positions or more (see Table 5). This suggests that the CRI ranking highlights aspects of countries' efforts to reduce 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 5: *Distribution of differences between pillars and CRI rankings*

Shifts with respect to CRI index	Spending pillar	Taxation pillar	Labour pillar
More than 30 positions	11 %	42 %	13 %
20 to 29 positions	15 %	20 %	19 %
10 to 19 positions	34 %	17 %	30 %
5 to 9 positions	14 %	11 %	20 %
Less than 5 positions	25 %	11 %	18 %
0 positions	2 %	1 %	1 %
Total	100 %	100 %	100 %
More than 10	61 %	78 %	62 %

Source: European Commission, Joint Research Centre (JRC), 2018.

The CRI index 2018 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 ([6]). 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. The pairwise correlation coefficients between the CRI index and the Gini and Palma measures are equal to −0.28 and −0.18, respectively. The correlation coefficients amount to −0.50 and −0.46 when the indices are computed only on the subsample of OECD countries (see Table 6). Along the same lines, Table 6 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 (0.56 vs 0.36), this being particularly true for OECD countries (0.77 vs 0.57). Though we should not interpret this as a causal relationship, it suggests that OECD countries with low levels of inequalities are those putting more efforts into ensuring a more equity-based society.

Table 6: *Comparison of the CRI index with other inequality-based rankings*

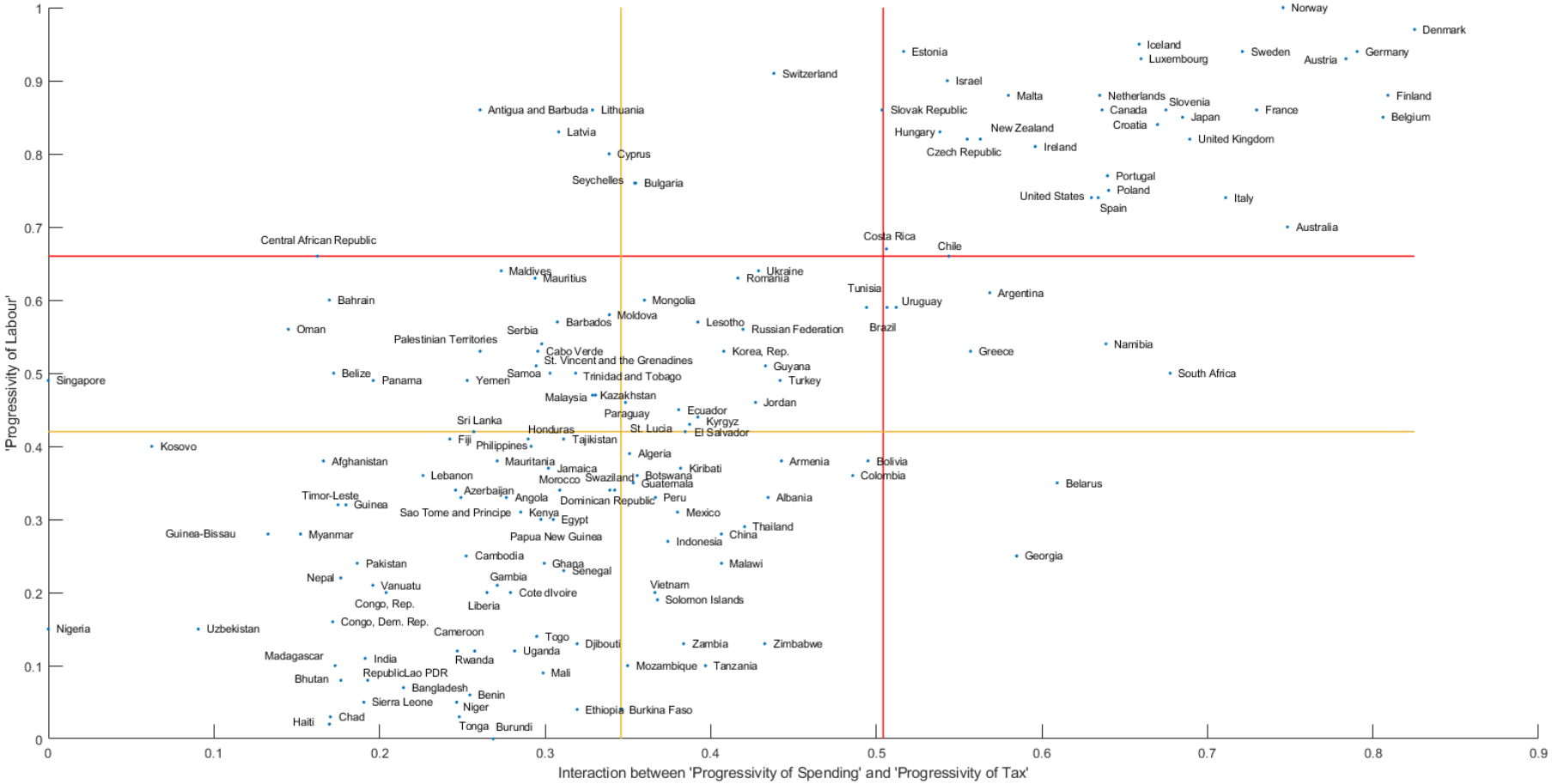
	CRI index	
Gini coefficient and Palma index	**Full sample**	**OECD countries**
Countries in the 75th percentile	0.36	0.57
Countries in the 25th percentile	0.56	0.77
Correlation coefficients		
Gini coefficient	−0.28	−0.50
Palma index	−0.18	−0.46

Source: European Commission, Joint Research Centre (JRC), 2018.

The CRI ranking and the ranking that results from the interaction of the spending and tax pillars differ by 10 positions or more for 54 % of the countries. This corroborates the finding that the CRI index is able to measure multidimensional aspects of the commitment to reducing inequality that cannot be captured by a single indicator. At this point, it is interesting to look further into the relationship between the two components that make up the CRI index: the labour pillar and the component that captures the interaction between the spending and tax pillars. Figure 6 depicts this relationship. The yellow lines in the plot represent the median values of the scores in each series across the countries included in the data set; the red lines represent the 75th percentiles. The analysis of this figure reveals some interesting findings. First, there is a linear positive relationship. In fact, the Pearson correlation coefficient is equal to 0.69 and statistically significant at 1 %. Second, 28 out of the 31 countries that are in the top 25 % of the best scores in both components (i.e. inside the top-right quadrant limited by the two red lines) belong to the OECD. This means that the most developed countries are those that have a strong commitment to reducing poverty through the labour market and the interaction between the progressivity of spending and tax. On the other side, the poorest countries such as Chad and Haiti implement poor policies to reduce inequality (see the bottom-left quadrant). Third, some countries present a mismatch in their policies. While countries have a strong commitment to reducing inequality through intervention in the labour market but not through public spending or taxes (e.g. countries inside of the upper-left quadrant such as Latvia and Lithuania); others show a high commitment in public spending and taxes, but not in adopting measures to distribute income through the labour market (e.g. countries in the bottom-right quadrant such as Belarus or Georgia).

([6]) The Gini coefficient and the data to calculate the Palma index can be freely downloaded from the World Development Indicators database at wdi.worldbank.org/table/1.3#.

Figure 6: *Relationship between the labour pillar and the interaction between the spending and tax pillars*



Source: European Commission Joint Research Centre, 2018.

NB: The interaction between progressivity of spending and progressivity of tax are represented in the x axis (geometric average of these pillars); while the progressivity of labour is represented in the y axis. Yellow lines represent median values across the countries included in the data set. Red lines represent 75th percentiles.

3. Impact of modelling assumptions on the Commitment to Reducing Inequality index

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 normalisation method, weighting scheme and aggregation formula, among other elements, all have an impact on CRI-based country ranks. These choices are based on expert opinion (e.g. selection of indicators and variables), common practice (e.g. min–max normalisation) 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.

The robustness assessment of the CRI index is based on the combination commonly used in the relevant literature on composite indicators, i.e. that of a Monte Carlo experiment and a multi-modelling approach (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 component of the index (i.e. the labour pillar and the interaction between the spending and tax pillars) and (c) the aggregation formula used to compute the overall CRI score ([7]). This type of uncertainty analysis aims to complement CRI country ranks with confidence intervals that help users of the index to appreciate for which countries ranks can be taken at face value and for which countries country ranks are to be analysed with caution because of their sensitivity to the methodological choices underlying the index computation.

The Monte Carlo simulations carried out related to the issue of weighting and comprised 1 000 runs. Each run corresponds to a different set of weights assigned to each of the two components of the index (i.e. the labour pillar and the interaction between the spending and tax pillars). The weights were randomly sampled from uniform continuous distributions centred at the weight value originally adopted for calculation of the CRI score. A perturbation of the weights ± 25 % around the reference values was adopted. For each simulation, weights were rescaled so that they always add up to one. The choice of the range for the weights' variation was driven by two different needs: ensuring a wide enough interval to have meaningful robustness checks and respecting the rationale of the CRI index calculation that is based on one third of the labour pillar and two thirds of the interaction effect of the spending and tax pillars. Given these considerations, limit values of uncertainty intervals for the component weights are 50-83 % for the interaction between the spending and tax pillars, and 25-42 % for the labour pillar (see Table 7).

For reasons of transparency and replicability, the CRI development team opted not to estimate the few missing values (see Table 1). 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 tax and labour pillars, and exclusively in non-OECD countries ([8]). To test the impact of this assumption, the JRC estimated missing values using the expectation maximisation (EM) algorithm ([9]).

([7]) 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-indicator-based rankings.

([8]) 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 (normalised) scores.

([9]) The 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

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' (Mariotti, 2018).

To capture this interaction, the developers 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 placing the three pillars — tax, spending and labour market — on equal footing whilst accounting for the interaction between tax and spending policies, each country's CRI score was computed as an arithmetic average made up by two thirds of the geometric average of the tax and spending pillars whilst 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 an 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 ([10]).

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 1 000 different sets of weight assigned to each of the three pillars. Table 7 summarises the uncertainties considered for the robustness assessment of the CRI index.

Table 7: *Uncertainty analysis for the CRI: weights, missing data, normalisation, aggregation*

	Reference	**Alternative**
I. Uncertainty in the treatment of missing values	No estimation of missing data	EM
II. Uncertainty intervals for the CRI component weights	Reference value for the weight	Distribution assigned for robustness analysis
Reducing inequality through the interaction between progressivity of spending and progressivity of tax $\left(\sqrt{\boldsymbol{spending} \cdot \boldsymbol{tax}}\right)$	0.66	U[0.50, 0.83]
Reducing inequality through labour rights and fair wages	0.33	U[0.25, 0.42]
III. Uncertainty in the aggregation formula at pillar level	$CRI\ 2018 = \frac{\sqrt{S \cdot T} + \sqrt{S \cdot T} + L}{3}$	Geometric average

Source: European Commission, Joint Research Centre (JRC), 2018.

parameter estimates; (2) the maximisation M-step: given a complete-data log likelihood, the M-step finds the parameter estimates to maximise the complete-data log likelihood from the E-step. The two steps are iterated until the iterations converge.

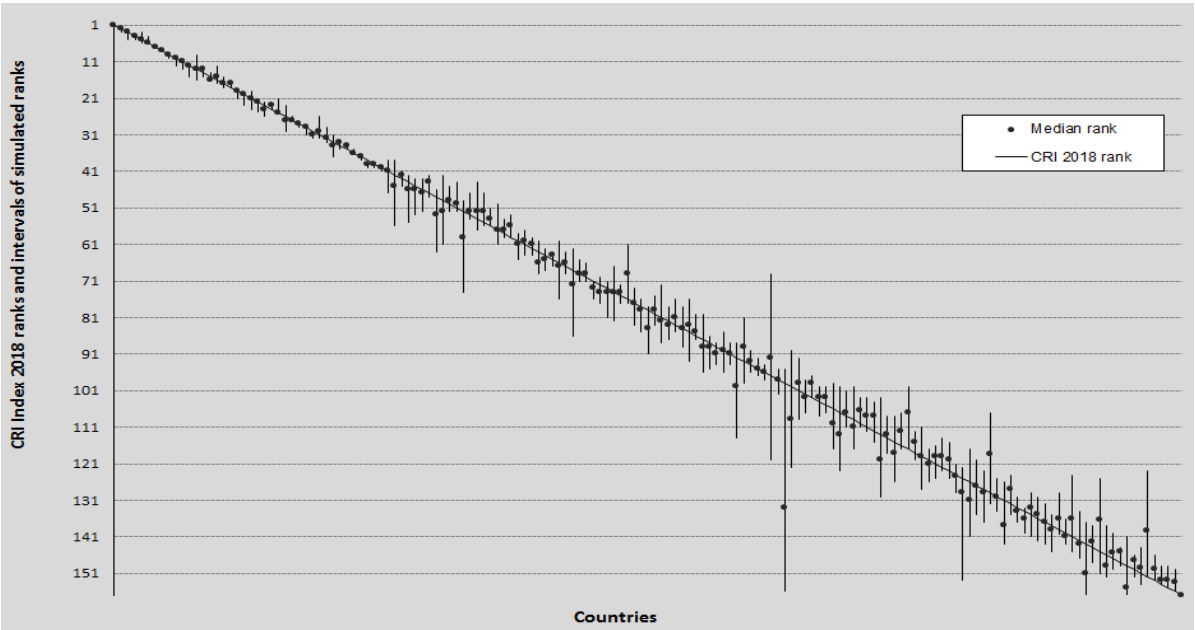
([10]) After renormalisation of the pillar scores, zero values were replaced with 0.00001 to avoid zero values in one pillar resulting in CRI scores equal to 0 regardless of the country's performance on the other two pillars.

3.1. Uncertainty analysis results

The robustness analysis results for the 157 countries are summarised in Figure 7, with median ranks and 90 % confidence intervals computed across the 4 000 Monte Carlo simulations of the overall CRI 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 8 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 index 2018 is a suitable summary measure.

CRI ranks are shown to be both representative of a plurality of scenarios and robust enough (for almost 30 % 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 five positions away) for 88 % of the countries suggests that the CRI index is a suitable summary measure. Furthermore, the narrow confidence intervals for 31 % of the countries' ranks (less than or equal to ± five 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 (e.g. Central African Republic) ([11]). For full transparency and information, Table 8 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. 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.

Figure 7: *Robustness analysis (CRI rank versus median rank, 90 % confidence intervals)*



Source: European Commission, Joint Research Centre (JRC), 2018.
NB: 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.

([11]) In fact, the Central African Republic has one of the highest levels of variability with an extremely wide confidence interval that covers 51 positions. This is not a surprising finding given that this country was detected as an outlier.

Table 8: *Country ranks and 90 % intervals for the CRI index 2018*

Country	Rank	Simulation		Country	Rank	Simulation		Country	Rank	Simulation	
		Median	Interval			Median	Interval			Median	Interval
Denmark	1	1	[1, 1]	Slovakia	30	31	[30, 32]	Jordan	59	56	[53, 59]
Germany	2	2	[2, 3]	South Africa	31	30	[26, 32]	Moldova	60	61	[58, 65]
Finland	3	3	[2, 5]	Namibia	32	32	[29, 33]	Armenia	61	60	[57, 64]
Austria	4	4	[4, 5]	Switzerland	33	34	[31, 37]	Kyrgyzstan	62	61	[59, 64]
Norway	5	5	[3, 6]	Argentina	34	33	[33, 35]	Mauritius	63	66	[60, 69]
Belgium	6	6	[4, 6]	Chile	35	34	[34, 35]	El Salvador	64	65	[62, 68]
Sweden	7	7	[7, 7]	Costa Rica	36	36	[35, 36]	Ecuador	65	64	[63, 67]
France	8	8	[8, 8]	Greece	37	37	[36, 38]	Albania	66	67	[60, 76]
Iceland	9	9	[9, 10]	Uruguay	38	39	[39, 40]	Saint Lucia	67	66	[63, 69]
Luxembourg	10	10	[10, 12]	Brazil	39	39	[39, 40]	Maldives	68	72	[62, 86]
Japan	11	11	[10, 13]	Tunisia	40	40	[40, 41]	Barbados	69	69	[65, 71]
Slovenia	12	12	[12, 15]	Belarus	41	41	[38, 47]	Paraguay	70	69	[66, 71]
Australia	13	13	[9, 16]	Lithuania	42	45	[38, 56]	Kazakhstan	71	73	[71, 76]
United Kingdom	14	13	[12, 15]	Ukraine	43	42	[41, 45]	Trinidad and Tobago	72	74	[70, 77]
Croatia	15	16	[14, 16]	Cyprus	44	46	[42, 55]	Serbia	73	74	[71, 81]
Italy	16	15	[12, 17]	Seychelles	45	46	[43, 53]	Thailand	74	74	[67, 82]
Netherlands	17	17	[15, 18]	Bulgaria	46	47	[43, 52]	Malaysia	75	74	[72, 77]
Canada	18	17	[16, 18]	Romania	47	44	[42, 48]	Kiribati	76	69	[61, 77]
Portugal	19	19	[19, 21]	Latvia	48	53	[46, 63]	Cape Verde	77	77	[73, 83]
Poland	20	20	[20, 23]	Georgia	49	52	[42, 61]	Samoa	78	79	[76, 85]
Malta	21	21	[19, 24]	Russia	50	49	[45, 52]	Saint Vincent and the Grenadines	79	84	[78, 91]
Spain	22	22	[21, 25]	Guyana	51	50	[44, 52]	Algeria	80	79	[75, 83]
United States	23	24	[22, 26]	Antigua and Barbuda	52	59	[49, 74]	China	81	82	[72, 88]
Ireland	24	23	[23, 24]	Turkey	53	52	[47, 54]	Peru	82	83	[78, 87]
Israel	25	25	[21, 25]	Bolivia	54	52	[44, 57]	Botswana	83	81	[76, 85]
Estonia	26	27	[23, 30]	Lesotho	55	52	[47, 56]	Mexico	84	84	[78, 89]
New Zealand	27	27	[26, 27]	South Korea	56	54	[51, 56]	occupied Palestinian territories	85	83	[76, 93]
Czechia	28	28	[27, 29]	Colombia	57	57	[50, 61]	Guatemala	86	85	[80, 87]
Hungary	29	29	[29, 31]	Mongolia	58	57	[54, 59]	Malawi	87	89	[80, 96]

Table 8: *Cont.*

Country	Rank	Simulation		Country	Rank	Simulation		Country	Rank	Simulation	
		Median	Interval			Median	Interval			Median	Median
Tajikistan	88	89	[86, 95]	Oman	113	120	[103, 130]	Myanmar/Burma	138	139	[135, 145]
Dominican Republic	89	91	[88, 94]	Ghana	114	113	[108, 118]	Nepal	139	136	[129, 140]
Indonesia	90	90	[85, 96]	Belize	115	118	[108, 126]	Benin	140	141	[136, 143]
Swaziland	91	91	[88, 94]	Azerbaijan	116	112	[107, 117]	Guinea-Bissau	141	136	[124, 145]
Zimbabwe	92	100	[88, 114]	São Tomé and Príncipe	117	107	[100, 117]	Niger	142	143	[134, 147]
Yemen	93	89	[81, 99]	Lebanon	118	115	[112, 120]	Burundi	143	151	[137, 157]
Philippines	94	93	[90, 96]	Mozambique	119	119	[111, 128]	Democratic Republic of the Congo	144	143	[138, 148]
Honduras	95	95	[92, 97]	Djibouti	120	121	[117, 126]	Tonga	145	137	[125, 151]
Jamaica	96	96	[94, 98]	Cambodia	121	119	[116, 125]	Kosovo	146	149	[138, 152]
Central African Republic	97	92	[69, 120]	The Gambia	122	119	[114, 123]	India	147	146	[140, 150]
Morocco	98	98	[95, 102]	Côte d'Ivoire	123	120	[115, 125]	Bangladesh	148	145	[144, 149]
Vietnam	99	133	[95, 156]	Liberia	124	125	[121, 129]	Singapore	149	155	[141, 157]
Bahrain	100	109	[90, 122]	Togo	125	129	[122, 153]	Laos	150	148	[146, 152]
Solomon Islands	101	99	[92, 109]	Burkina Faso	126	131	[117, 141]	Madagascar	151	150	[144, 154]
Sri Lanka	102	103	[98, 107]	Afghanistan	127	127	[120, 135]	Bhutan	152	140	[123, 152]
Mauritania	103	99	[97, 103]	Mali	128	129	[123, 137]	Sierra Leone	153	150	[146, 153]
Egypt	104	103	[100, 108]	Guinea	129	119	[107, 132]	Chad	154	153	[149, 154]
Papua New Guinea	105	103	[100, 107]	Uganda	130	130	[125, 134]	Haiti	155	153	[149, 155]
Zambia	106	110	[99, 117]	Ethiopia	131	138	[126, 143]	Uzbekistan	156	154	[150, 156]
Tanzania	107	113	[100, 123]	Timor-Leste	132	128	[124, 135]	Nigeria	157	157	[156, 157]
Fiji	108	107	[101, 111]	Rwanda	133	134	[130, 137]				
Panama	109	111	[100, 117]	Cameroon	134	136	[133, 140]				
Kenya	110	107	[103, 111]	Congo	135	133	[129, 141]				
Angola	111	108	[103, 112]	Vanuatu	136	135	[130, 142]				
Senegal	112	108	[104, 115]	Pakistan	137	137	[132, 143]				

Source: European Commission, Joint Research Centre (JRC), 2018.

NB: 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.

3.2. Sensitivity analysis results

To complement to the uncertainty analysis, the sensitivity analysis has been used to identify which of the modelling assumptions has the highest impact on country ranks. Table 9 summarises 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).

When the geometric average is used across all three pillars, there are nine countries that decline between 10 and 19 positions — Antigua and Barbuda, Bahrain, Burkina Faso, Burundi, Central African Republic, Ethiopia, Oman, Tanzania and Zimbabwe. No country improves by 10 or more positions. The impact of estimating missing data causes two countries — Central African Republic and Bhutan — to improve by 20 positions or more; and six countries improve between 10 and 19 positions — Guinea, Guinea-Bissau, Kiribati, São Tomé and Príncipe, Tonga, and Yemen. On the other hand, two countries deteriorate by 10 or more positions (Togo and Vietnam). 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 since 20 countries improve or deteriorate by 10 or more positions. 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).

Table 9: *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 and 19 positions**		**by 20 or more positions**		**between 10 and 19 positions**	
Geometric average versus original CRI aggregation formula (pillar level)	0		0		0		9	Antigua and Barbuda, Bahrain, Tanzania, Burkina Faso, Ethiopia, Burundi, Zimbabwe, Oman, Central African Republic.
EM imputation vs no imputation of missing data (13-indicator data set)	2	Central African Republic, Bhutan	6	Kiribati, São Tomé and Príncipe, Guinea, Guinea-Bissau, Tonga, Yemen.	1	Vietnam	1	Togo
Geometric average and EM imputation vs original CRI aggregation formula and no estimation of missing values	2	Guinea, Bhutan	7	Central African Republic, Kiribati, Lesotho, São Tomé and Príncipe, Guinea-Bissau, Tonga, Yemen.	2	Togo, Vietnam	9	Antigua and Barbuda, Bahrain, Burkina Faso, Ethiopia, Burundi, Zimbabwe, Tanzania, Oman, Latvia.

Source: European Commission, Joint Research Centre (JRC), 2018.

4. Conclusion

Development Finance International and Oxfam have developed the CRI index with a view to monitoring the extent to which governments are tackling the gap between rich and poor through three policy areas: public spending, taxes and labour market. Overall, the CRI framework is well constructed and a lot of thought has clearly been put into it. One of its greatest strengths is the amount of original research into the policies for fighting inequality and the transparency and detail of all data associated with the index, as well as the extensive documentation on the methodology and the online tool accompanying the index.

The JRC statistical audit has delved into the workings of the CRI index framework to assess the statistical properties of the data and the methodology used in the index construction. The key findings of the statistical assessment conducted herein are the following.

First, the results offer statistical justification for the theoretical framework underpinning the CRI index. The conceptual grouping of the nine indicators into three pillars is statistically confirmed. The aggregation of the spending and tax pillars into one component to capture the interaction of these two policies also seems adequate.

Second, the JRC analysis suggests that the conceptualised multilevel structure of the CRI index is statistically coherent and balanced (i.e. not dominated by any pillar, and all nine indicators contributing, to a greater or lesser extent, to the variation of their respective pillar scores).

Third, the CRI country ranks are robust regarding methodological assumptions related to the estimation of missing data, weighting and aggregation formula. It is reassuring that for 88 % of the 157 countries, the CRI rank is close (less than five 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 almost one third of the countries, the confidence intervals are narrow enough to allow for inferences to be drawn (less than or equal to ± five positions). 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.

Fourth, one way in which the CRI helps to highlight what governments in 157 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 overall CRI index and each of the three pillars: for more 61 % (up to 76 %) of the 157 countries included in the CRI index 2018, the CRI ranking and any of the three pillar rankings differ by 10 positions or more. This outcome both evidences 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 input for evidence-informed policymaking.

Fifth, relevant insights may emerge when analysing the relationship between the labour pillar and the component that captures the interaction between the spending and tax pillars. It is detected that the most developed countries are those that have a strong commitment to reducing poverty through the labour market and the interaction of the progressivity of spending and tax.

Sixth, some points that call for possible refinements of the CRI framework have also been identified.

- The data set must be previously checked in order to detect possible outliers that could distort the score and rank of the index. This is the case for the value of the Central African Republic in the minimum wage indicator.

- The social spending and tax structure indicators should be normalised to cover the whole interval [0, 1].
- Three out of the nine indicators, namely harmful tax practices and tax structure under the tax pillar and minimum wage under the labour pillar, were found not to be influential at the overall CRI level, though they were at their own respective pillar levels. This is indicative of a different behaviour of these indicators with respect to the remaining ones composing the CRI index. The JRC recommends that the CRI developing team keep these indicators in the current framework because of their conceptual relevance to the phenomenon, but test and eventually refine next year's release along these issues if next year's data confirm the same pattern.

All things considered, the present audit findings confirm that the Commitment to Reducing Inequality Index 2018 meets international quality standards for statistical soundness. 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. Having said this, it is worth mentioning that the CRI index cannot be understood as a substitute of context-specific knowledge of each country's path to reducing inequality, or as a substitute for a detailed analysis of each government's proposals or positions. However, it is a useful tool that should enable policymakers to derive more accurate and meaningful conclusions, and to potentially guide choices on priority setting and policy formulation.

References and related reading

Becker, W., Saisana, M., Paruolo, P. and Vandecasteele, I., 2017, 'Weights and importance in composite indicators: Closing the gap', *Ecological Indicators*, Vol. 80, pp. 12-22.

Groeneveld, R.A. and Meeden, G., 1984, 'Measuring skewness and kurtosis' *The Statistician*, Vol. 33, pp. 391-399.

Little, R.J.A., Rubin, D.B., 2002, *Statistical Analysis with missing data, second edition*, John Wiley & Sons, New York.

Mariotti, C., 2018, The commitment to reducing inequality index 2018. Methodology. In M. Lawson & M. Martin (2018). The Commitment to Reducing Inequality Index 2018: A global ranking of governments based on what they are doing to tackle the gap between rich and poor. Retrieved from https://www.oxfamamerica.org/static/media/files/The_Commitment_to_Reducing_Inequality_Index_2018_Methodology.pdf

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

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

OECD/EC JRC, 2008, *Handbook on constructing composite indicators: Methodology and user guide*, OECD, Paris.

Paruolo, P., Saisana, M., Saltelli, A., 2013, 'Ratings and rankings: Voodoo or Science?', *Journal of the Royal Statistical Society A*, Vol. 176, No 3, pp. 603-634.

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

Saisana, M., Saltelli, A., 2011, 'Rankings and ratings: Instructions for use', *Hague Journal on the Rule of Law*, Vol. 3, No 2, pp. 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*, Vol. 168, No 2, pp. 307-323.

Saltelli, A., Funtowicz, S. 2014. 'When all models are wrong', *Issues in Science and Technology*, No 2 Winter 2014, pp. 79-85.

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

List of figures

List of tables

KJ-NA-29406-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:10.2760/942703

ISBN 978-92-79-96744-3