APPENDIX IV

ADJUSTMENTS TO THE GLOBAL INNOVATION INDEX FRAMEWORK, YEAR-ON-YEAR COMPARABILITY OF RESULTS, AND TECHNICAL NOTES

Adjustments to the Global Innovation Index framework

The Global Innovation Index (GII) is a cross-economy performance assessment, compiled on an annual basis, which continuously seeks to update and improve the way innovation is measured. The GII report pays special attention to making the statistics used in the Economy Profiles and Data Tables accessible by providing data sources and definitions, and detailing the computation methodology (Appendix II, III, and IV). This Appendix summarizes the changes made this year and provides an assessment of the impact these changes have on the comparability of rankings.

Beyond the use of the World Intellectual Property Organization (WIPO) data, we collaborate with public international bodies, such as the International Energy Agency, the United Nations Educational, Scientific and Cultural Organization (UNESCO), the United Nations Industrial Development Organization (UNIDO), the World Trade Organization (WTO), and the Joint Research Centre of the European Commission (JRC). We also collaborate with private organizations, such as Brand Finance, IHS Markit, ZookNIC Inc, Thomson Reuters, Wikimedia Foundation, and AppAnnie to obtain the best globally available data on innovation.

Table A-IV.1 provides a summary of adjustments to the GII 2020 framework. A total of 10 indicators were modified this year. Six indicators had a methodology change at source, two underwent methodological changes, one new indicator was replaced, and one changed in code only.

Methodology and data

The methodologies for computing indicators 3.3.2, 3.3.3, 4.2.1, 4.2.3, 6.2.4, and 7.2.3 were all changed by the corresponding data source institutions; therefore, the scores calculated under the old methodology are not comparable to the new scores. The methodology underpinning indicator 5.2.3 GERD financed by abroad was updated. This year, the indicator is scaled by current GDP rather than as a percentage of total gross

TABLE A-IV.1

Changes to the GII 2020 framework

GII 2019	Adjustment	GII 2020
3.3.2 Environmental performance	Indicator changed at source	3.3.2 Environmental performance
3.3.3 ISO 140001 environxment certificates/bn PPP$ GDP	Indicator changed at source	3.3.3 ISO 140001 environment certificates/bn PPP$ GDP
4.2.1 Ease of protecting minority investors	Indicator changed at source	4.2.1 Ease of protecting minority investors
4.2.3 Venture capital deals/bn PPP$ GDP	Indicator changed at source	4.2.3 Venture capital deals/bn PPP$ GDP
5.2.3 GERD financed by abroad, %	Methodology changed	5.2.3 GERD financed by abroad, % GDP
6.2.4 ISO 9001 quality certificates/bn PPP$ GDP	Indicator changed at source	6.2.4 ISO 9001 quality certificates/bn PPP$ GDP
7.1.2 Industrial designs by origin/bn PPP$ GDP	Code changed	7.1.3 Industrial designs by origin/bn PPP$ GDP
7.1.3 ICTs & business model creation	Replaced	7.1.2 Global brand value / bn PPP$ GDP
7.2.1 Cultural & creative services exports, % total trade	Methodology changed	7.2.1 Cultural & creative services exports, % total trade
7.2.3 Entertainment & Media market/th pop. 15-69	Indicator changed at source	7.2.3 Entertainment & Media market/th pop. 15-69

Source: Global Innovation Database, Cornell, INSEAD, and WIPO.
Notes: Refer to Appendix I and III for a detailed explanation of terminology and acronyms. Refer to Appendix III for a detailed explanation of methodological changes at source.

expenditure on R&D in a country. Indicator 7.2.1 Cultural & creative services exports is based on a revised list of cultural and creative services, to align closer to the 2009 UNESCO Framework on Cultural Statistics.[1] The services included are: information services (EBOPS 2010 code SI3); Advertising, market research, and public opinion polling services (SJ22); Audio-visual and related services (SK1); and Heritage and recreational services (SK23) (Appendix III). One new indicator 7.1.2 Global brand value, top 5,000 was added to the model replacing indicator 7.1.3 on ICT and business model creation. This also resulted in a change in the numbering of one 2019 indicator 7.1.2, that is now labeled 7.1.3.

Missing values

Since its inception, one of the core missions of the GII is to increase awareness of the importance of collecting data. The GII has helped to improve the number of data points submitted to international data agencies. In the GII 2020, with the inclusion of four economies in the GII sample, coverage remains relatively close to the level seen last year, with 10% of data points missing.

When it comes to economy coverage, the objective is to include as many as possible. However, it is also important to maintain a good level of data coverage within each of these economies. Because the GII results depend on data availability (Appendix V), which in turn affects the overall GII rankings, the threshold rule for economies with missing data and the minimum coverage necessary per sub-pillar were progressively tightened in 2016 and 2017 (Appendix IV: Technical Notes).

The motivation behind the introduction of these adjustments is because of data availability, which, historically, was less satisfactory when considering innovation outputs in the GII. For instance, this year, 18.8% of all economies show data coverage of less than 75% but exhibit over 66% coverage in the Output Sub-Index, while only 1.3% of these economies have this coverage range in the Input Sub-Index. This year, four new economies, Cabo Verde, the Lao People's Democratic Republic, Myanmar, and Uzbekistan are included in the GII 2020 due to data coverage improving to above the 66% threshold in the Output Sub-Index. Conversely, Burundi and Nicaragua drop from the GII economy sample due to data coverage being below the 66% threshold in the Output Sub-Index.

Despite the requirement for a minimum level of coverage, for several economies the number of missing data points remains high. Table A-IV.2 lists the economies with the highest number of missing data points (20 or more).

Conversely, Table A-IV.3 lists economies with the best data coverage. These economies are missing five data points at the most, while others are missing none.

For the last three years, more stringent rules were introduced, resulting in significant data coverage improvements for various economies. Table A-IV.4 shows economies with improved data coverage from 2016 to 2020. The list compiles those economies that have consistently improved the number of indicators with data available from year to year during that time period. At the same time, fewer economies had a decline in data coverage, as shown in Table A-IV.5. In particular, Uzbekistan, which is a new addition to the GII sample, displayed a noteworthy improvement in the coverage of its Output Sub-Index variables showing this year's data for 8 additional indicators, when compared to the data collected in 2019.

Year-on-year comparability of results—sources of change in the rankings

The GII compares the performance of national innovation systems across economies and presents the changes in economy rankings over time.

Importantly, scores and rankings from one year to the next are not directly comparable (see GII 2013, Annex 2, for a full explanation). Making inferences about absolute or relative performance based on year-on-year differences in rankings can be misleading. Each ranking reflects the relative positioning of a particular economy based on the conceptual framework, data coverage, and the sample of economies in a given year, also reflecting changes in the underlying indicators at source and in data availability.

A few factors influence year-on-year rankings of an economy:

- the actual performance of the economy in question;
- adjustments made to the GII framework;
- data updates, the treatment of outliers, and missing values; and
- the inclusion or exclusion of economies in the sample.

Additionally, the following characteristics complicate the time-series analysis based on simple GII scores or rankings:

- **Missing values.** The GII produces relative index scores, which means that a missing value for one economy affects the index score of other economies. Because the number of missing values decreases every year, this problem reduces over time.
- **Reference year.** The data underlying the GII do not refer to a single year but to several years, depending on the latest available year for any given variable. In addition, the reference years for different variables are not the same for each economy. The motivation for this approach is that it widens the set of data points for cross-economy comparability.
- **Normalization factor.** Most GII variables are normalized using either GDP or population, with the intention to enable cross-economy comparability. Yet, this implies that year-on-year changes in individual variables may be driven either by the variable's numerator or by its denominator.
- **Consistent data collection.** Measuring the change of year-on-year performance relies on the consistent collection of data over time. Changes in the definition of variables or in the data collection process could create movements in the rankings that are unrelated to performance.

TABLE A-IV.2

GII economies with the most missing values

Economy	Number of missing values
Guinea	23
Cabo Verde	22
Lao People's Democratic Republic	21
Niger	20

Source: Global Innovation Database, Cornell, INSEAD, and WIPO.

TABLE A-IV.3

GII economies with the fewest missing values

Economy	Number of missing values	Economy	Number of missing values	Economy	Number of missing values
Chile	0	Russian Federation	2	Argentina	4
Indonesia	0	Slovakia	2	Belgium	4
Malaysia	0	Spain	2	Canada	4
Mexico	0	Brazil	3	Ireland	4
Thailand	0	Costa Rica	3	Kenya	4
Colombia	1	Cyprus	3	Latvia	4
Hungary	1	Denmark	3	Lithuania	4
Poland	1	Estonia	3	Luxembourg	4
Portugal	1	Finland	3	New Zealand	4
Romania	1	France	3	Republic of Moldova	4
Ukraine	1	India	3	South Africa	4
Austria	2	Israel	3	Turkey	4
Bulgaria	2	Kazakhstan	3	United Kingdom	4
Czech Republic	2	Malta	3	United States of America	4
Germany	2	Netherlands	3	Uruguay	4
Greece	2	Norway	3	Viet Nam	4
Italy	2	Serbia	3	Australia	5
Morocco	2	Singapore	3	Croatia	5
Philippines	2	Slovenia	3	Georgia	5
Republic of Korea	2	Sweden	3	Panama	5
		Switzerland	3	Tunisia	5

Source: Global Innovation Database, Cornell, INSEAD, and WIPO.

TABLE A-IV.4

Indicator coverage improvement, from 2016 to 2020, in % and number

Economy	2016-2020	Improvement	Number
Yemen	from 29 to 18	11.24%	11
Algeria	from 17 to 7	19.89%	10
Honduras	from 17 to 7	14.93%	10
United Arab Emirates	from 21 to 11	19.89%	10
Bahrain	from 17 to 8	15.91%	9
Cambodia	from 18 to 9	13.88%	9
Jordan	from 20 to 11	17.18%	9
Mozambique	from 20 to 11	13.88%	9
Côte d'Ivoire	from 16 to 8	10.13%	8
Iran (Islamic Republic of)	from 18 to 10	15.91%	8
Montenegro	from 22 to 14	13.67%	8
Rwanda	from 22 to 14	10.68%	8
Tajikistan	from 23 to 15	10.68%	8
Belarus	from 10 to 3	14.54%	7
Benin	from 15 to 8	8.26%	7
Burkina Faso	from 18 to 11	8.67%	7
Malta	from 23 to 16	25.99%	7
Namibia	from 24 to 17	11.58%	7
Niger	from 26 to 19	7.23%	7
Togo	from 27 to 20	7.54%	7
Mali	from 8 to 2	9.64%	6
Morocco	from 10 to 4	29.29%	6
Viet Nam	from 18 to 12	20.47%	6
Ethiopia	from 7 to 2	8.94%	5
Ghana	from 8 to 3	8.94%	5
Jamaica	from 9 to 4	6.24%	5
Kenya	from 14 to 9	18.35%	5
Netherlands	from 16 to 11	21.75%	5
Oman	from 16 to 11	10.46%	5
Spain	from 22 to 17	26.89%	5

Source: Global Innovation Database, Cornell, INSEAD, and WIPO.
Notes: Annualized growth.

TABLE A-IV.5

Indicator coverage decline, from 2016 to 2020, in % and number

Economy	2016-2020	Improvement	Number
Japan	from 2 to 6	31.61%	4
Uganda	from 13 to 16	5.33%	3
Australia	from 3 to 5	13.62%	2
Bolivia (Plurinational State of)	from 12 to 14	3.93%	2
Madagascar	from 15 to 17	3.18%	2
South Africa	from 2 to 4	18.92%	2
Turkey	from 2 to 4	18.92%	2

Source: Global Innovation Database, Cornell, INSEAD, and WIPO.
Note: Annualized growth.

A detailed economy study based on the GII database and the economy profile over time, coupled with analytical work on the ground, including innovation actors and decision-makers, yields the best results in terms of grasping an economy's innovation performance over time as well as in identifying possible avenues for improvement.

Technical notes

Audit by the European Commission's Competence Centre on Composite Indicators and Scoreboards (COIN) at the Joint Research Centre (JRC)

The JRC-COIN has extensively researched the complexity of composite indicators that rank economies' performances along policy lines. For the tenth consecutive year, the JRC-COIN has performed a thorough "robustness" and "sensitivity" analysis of the GII to assess structural changes that are made to the list of indicators by the GII team (Table A-IV.1).

The recommendations from the JRC-COIN audit on the GII 2019 model were reviewed and incorporated into the GII 2020 model. This year, for an economy to feature in the GII 2020, the minimum symmetric data coverage is at least 35 indicators in the Innovation Input Sub-Index (66%) and 18 indicators in the Innovation Output Sub-Index (66%), with scores for at least two sub-pillars per pillar. In 2020, consideration was given to whether scores for all sub-pillars, for all pillars, would be required for economies to be considered in the GII. Ultimately, this rule was not applied this year but will be reviewed again in 2021 and implemented if applicable.

A final audit of the GII 2020 model was performed in May 2020 (Appendix V).

Composite indicators

The GII relies on seven pillars, each divided into three sub-pillars, of which each include two to five individual indicators. Sub-pillar scores are calculated using the weighted average of its individual indicators. Pillar scores are calculated using the weighted average of its sub-pillar scores.

The notion of weights as important coefficients was revised this year in more detail to ensure a greater statistical coherence of the model, following the recommendations of the JRC-COIN.[2]

The GII includes three indices:

1. The Innovation Input Sub-Index is the average of the first five pillar scores.
2. The Innovation Output Sub-Index is the average of the last two pillar scores.
3. The Global Innovation Index is the average of the Input and Output Sub-Indices.

Economy rankings are provided for indicators, sub-pillars, pillars, and index scores.

Individual indicators

The GII 2020 model includes 80 indicators, which fall into three categories:

1. quantitative/objective/hard data (58 indicators),
2. composite indicators/index data (18 indicators), and
3. survey/qualitative/subjective/soft data (4 indicators).

Hard data

Hard data (58 indicators) are drawn from a variety of public and private sources. These include, among others, the United Nations Educational, Scientific and Cultural Organization (UNESCO), the United Nations Industrial Development Organization (UNIDO), the World Intellectual Property Organization (WIPO), the World Bank, the Joint Research Centre of the European Commission (JRC), PwC, Thomson Reuters, IHS Markit, Wikimedia Foundation, and AppAnnie.

This year an indicator showing which economies have the most valuable brands based on Brand Finance data is introduced. This indicator assesses the economy's brands in the top 5,000 global brand database and produces the sum of the brand values corresponding to that economy. This sum is then scaled by GDP.[3]

Indicators are often correlated with population, GDP, or some other size-related factor; they require scaling by a relevant size indicator for economy comparisons to be valid. Most indicators are either scaled at source or do not need to be scaled; for the rest, the scaling factor was chosen to represent a fair picture of economy differences. Scaling affected 42 indicators, which can be broadly divided into four groups:

1. Indicators 2.1.1, 2.3.2, 3.2.3, 4.1.2, 4.1.3, 4.2.2, 5.1.3, 5.2.3, 5.3.4, 6.2.3, and 6.3.4 are scaled by GDP in current US$.[4]

2. Indicators 3.3.3, 4.2.3, 5.2.4, 5.2.5, 6.1.1, 6.1.2, 6.1.3, 6.1.4, 6.2.4, 7.1.1, 7.1.2, 7.1.3, and 7.3.4 are scaled by GDP in purchasing power parity current international dollars. This choice of denominator was dictated by a willingness to appropriately account for differences in development stages; in addition, scaling these variables by population would improperly bias results to the detriment of economies with a large young or aging population.[5]

3. Indicators 3.2.1, 5.1.5, 6.2.2, 7.2.2, 7.2.3, 7.3.1, 7.3.2, and 7.3.3 are scaled by population. Total population for 3.2.1, population 25+ years old for 5.1.5, population 15–64 years old for 6.2.2, and population 15–69 years old for the remaining.[6]

4. Indicators 5.3.1, 5.3.2, 5.3.3, 6.3.1, 6.3.2, 6.3.3, 7.2.1, and 7.2.5 are scaled by total trade; and indicators 6.2.5 and 7.2.4 by the total unit used to measure the particular statistic.[7]

Indices

Composite indicators are collected from a series of specialized agencies and academic institutions, such as the World Bank, the UN Public Administration Network (UNPAN), and Yale and Columbia Universities. Statisticians discourage the use of an "index within an index" on two main grounds: the distorting effect of the different computing methodologies used and the risk of duplicating variables. The normalization procedure partially solves the former (more on this below). To avoid the mistake of including a particular indicator more than once (directly and indirectly through a composite indicator), only indices with a narrow focus (18 in total) were selected.

Any additional disadvantage is outweighed by what is gained with model parsimony, acknowledgment of expert opinion, and focus on multi-dimensional phenomena that can hardly be captured by a single indicator.[8]

Survey data

Survey data are drawn from the World Economic Forum's Executive Opinion Survey (EOS). Survey questions are drafted to capture subjective perceptions on specific topics. Four EOS questions were retained to capture phenomena strongly linked to innovative activities for which hard data are nonexistent or have low coverage for economies.

Economy coverage and missing data

This year the GII covers 131 economies, selected based on the availability of data, and achieves the same percentage of indicator coverage as in the GII 2019 (Appendix IV: Technical Notes).

For each economy, only the most recent yearly data was considered. As a rule, the GII enforced the cut-off year to be 2010 for considering data at the indicator level. A few exceptions were made for years prior to the cut-off year.[9] For the sake of transparency and replicability of results, no additional effort was made to fill missing values. Missing values are indicated with "n/a" and are not considered in the sub-pillar score. However, the JRC-COIN audit assessed the robustness of the GII modeling choices (i.e., no imputation of missing data, fixed predefined weights, and arithmetic averages) by imputing missing data, applying random weights, and using geometric averages. Since 2012, based on this assessment, a confidence interval has been provided for each ranking in the GII as well as the Input and Output Sub-Indices (Appendix V).

Treatment of series with outliers

Potentially problematic indicators with outliers that could polarize results and unduly bias the rankings were treated according to the rules listed below, as per the recommendations of the JRC-COIN. This affected 29 indicators; 28 out of the 58 hard data indicators and 1 out of the 18 composite indicators.

First rule: selection

Problematic indicators were identified by skewness or kurtosis. The problematic indicators had either:

- an absolute value of skewness greater than 2.25, or
- a kurtosis greater than 3.5.[10]

Second rule: treatment

Series with one to five outliers (24 cases) were winsorized; the values distorting the indicator distribution were assigned the next highest value, up to the level where skewness and/or kurtosis entered within the ranges specified above.[11]

Series with five or more outliers, skewness and/or kurtosis entered within the ranges specified above after multiplication by a given factor *f* and transformation by natural logs.[12] Since only "goods" were affected (i.e., indicators for which higher values indicate better outcomes, as opposed to "bads"), the formula used was:

$$\ln\left[\frac{(Max \times f - 1)(economy\ value - Min)}{Max - Min} + 1\right]$$[13]

where "min" and "max" are the minimum and maximum indicator sample values.

Normalization

The 80 indicators were then normalized into the [0, 100] range, with higher scores representing better outcomes. Normalization was according to the min-max method; where the min and max values were given by the minimum and maximum indicator sample values, respectively. The exception for index and survey data, for which the original series range of values was kept as min and max values (for example, [0, 1] for UNPAN indices; [1, 7] for the World Economic Forum Executive Opinion Survey questions; [0, 100] for World Bank's World Governance Indicators; etc.). The following formula was applied:

$$\text{Goods:}\ \frac{economy\ value - Min}{Max - Min} \times 100$$

$$\text{Bads:}\ \frac{Max - economy\ value}{Max - Min} \times 100$$

Notes:

1 See: http://uis.unesco.org/sites/default/files/documents/unesco-framework-for-cultural-statistics-2009-en_0.pdf

2 Paruolo et al., 2013 show that a theoretical inconsistency exists between the real theoretical meaning of weights and the meaning generally attributed to them by the standard practice in constructing composite indicators that use them as importance coefficients in combination with linear aggregation rules. The approach followed in the GII this year, as last year, is to assign weights of 0.5 or 1 to each component in a composite to ensure the highest correlations between them (i.e., indicator/sub-pillar, sub-pillar/pillar, etc.). Two sub-pillars (7.2 Creative goods and services, and 7.3 Online creativity) and 27 indicators (1.1.1, 1.2.1, 1.2.2, 2.1.4, 2.1.5, 2.2.3, 3.1.1, 3.1.2, 3.1.3, 3.1.4, 3.2.1, 3.2.2, 4.1.3, 4.3.1, 5.1.5, 5.2.1, 5.3.3, 5.3.4, 6.1.2, 6.2.2, 6.3.1, 6.3.4, 7.1.3, 7.2.1, 7.2.2, 7.2.3, and 7.2.4) are weighted 0.5; the rest have a weight of 1. This year the weights for 24 indicators were adjusted to provide higher statistical coherence (3.1.1 ICT access, 3.1.2 ICT use, 3.1.3 Government's online service, 3.1.4 E-participation, 4.1.3 Microfinance gross loans, 5.3.3 ICT services imports, 5.3.4 FDI net inflows, and 7.2.4 Printing & other media now have a weight of 0.5; 2.2.1 Tertiary enrolment, 3.3.3 ISO 14001 environmental certificates, 4.2.2 Market capitalization, 4.2.3 Venture capital deals, 4.3.2 Intensity of local competition, 5.1.3 GERD performed by business, 5.1.4 GERD financed by business, 5.2.4 JV–strategic alliance deals, 5.3.1 Intellectual property payments, 6.1.1 Patents by origin, 6.1.4 Scientific & technical articles, 6.2.3 Computer software spending, 6.2.4 ISO 9001 quality certificates, 6.2.5 High- & medium-high-tech manufacturing, 6.3.2 High-tech net exports, and 6.3.3 ICT services exports now have a full weight of 1).

3 Brand Finance's study is based on publicly available information on the largest brands in the world. Brand Finance calculates brand value using the Royalty Relief methodology, which determines the value a company would be willing to pay to license its brand as if it did not own it. The methodology is compliant with industry standards set in ISO 10668. This approach involves estimating the future revenue attributable to a brand and calculating a royalty rate that would be charged for the use of the brand.

4 These indicators are Expenditure on education (2.1.1); Gross expenditure on R&D (GERD) (2.3.2); Gross capital formation (3.2.3); Domestic credit to private sector (4.1.2); Microfinance institutions' gross loan portfolio (4.1.3); Market capitalization (4.2.2); GERD performed by business enterprise (5.1.3); GERD financed by abroad (5.2.3); Foreign direct investment net inflows (5.3.4); Total computer software spending (6.2.3); and Foreign direct investment net outflows (6.3.4).

5 These count variables are mainly indicators that increase disproportionately with economic growth. They include: ISO 14001 environmental certificates (3.3.3); Venture capital deals; (4.2.3) joint venture/strategic alliance deals; (5.2.4) Patent families filed in two or more offices (5.2.5); Patent applications by origin (6.1.1); PCT international applications by origin (6.1.2); Utility model applications by origin (6.1.3); Scientific and technical publications (6.1.4); ISO 9001 quality certificates (6.2.4); Trademark application class count by origin (7.1.1); Global brand value, top 5000 (7.1.2), Industrial designs by origin (7.1.3); and Mobile app creation (7.3.4).

6 These variables are Electricity output (3.2.1); Females employed with advanced degrees (5.1.5); New business density (6.2.2); National feature films produced (7.2.2); Entertainment and media market (7.2.3); Generic (7.3.1) and Country-code (7.3.2) top-level Internet domains; and Wikipedia yearly edits (7.3.3).

7 Intellectual property payments (5.3.1); High-tech net imports (5.3.2); ICT services imports (5.3.3); Intellectual property receipts (6.3.1); High-tech net exports (6.3.2); ICT services exports (6.3.3); Cultural and creative services exports (7.2.1); and Creative goods exports (7.2.5) were scaled by total trade; High-tech and medium-high-tech output (6.2.5) and Printing and other media (7.2.4) were scaled by total manufacturing output.

8 For example, GII sub-pillar 3.1 Information and communication technologies (ICTs) is composed of four indices: ICT Access and Use sub-indices, and UNPAN's Government Online Service and E-Participation indices. The first two, previously part of ITU's ICT Development Index, are now produced by the GII independently from other components from that original index, following the methodology of the ITU's ICT Development Index 2017. Similarly, the Online Service Index is a component of UNPAN's E-Government Development Index together with two indices on Telecommunication Infrastructure and Human Capital that were not considered, as they duplicate GII pillars 3 and 2, respectively. The e-Participation Index was developed separately by UNPAN in 2010.

9 A total of 18 economies in 11 indicators show data that is previous to 2010. These are Botswana (2009), Philippines (2009), and Morocco (2009), in Expenditure on education (2.1.1); Botswana (2009) and Qatar (2009) in Government funding per pupil (2.1.2); Kenya (2009) in School life expectancy (2.1.3); Israel (2009), Kenya (2009), and Trinidad and Tobago (2009) in Pupil-teacher ratio (2.1.5); Bangladesh (2009) in Tertiary inbound mobility (2.2.3); the Plurinational State of Bolivia (2009) in Gross expenditure on R&D (2.3.2); Botswana (2009), Chile (2009), Costa Rica (2009), Jamaica (2009), Mexico (2009), Panama (2009), and Trinidad and Tobago (2009) in Firms offering formal training (5.1.2); the Plurinational State of Bolivia (2009) and Burkina Faso (2009) in GERD financed by business (5.1.4); Tajikistan (2009) in Females employed with advanced degrees (5.1.5); Malawi (2009) in New businesses (6.2.2); and Cameroon (2009) in National feature films (7.2.2.).

10 Based on Groeneveld and Meeden (1984), which sets the criteria of absolute skewness above 1 and kurtosis above 3.5. The skewness criterion was relaxed to account for the small sample at hand (131 economies).

11 This distributional issue affects the following variables: 2.1.5, 3.2.1, 5.2.3, 5.3.2, 6.1.5, 7.1.1, 7.2.2, and 7.2.4, (1 outlier); 4.2.2, 4.2.3, 5.3.1, 6.1.3, 7.1.3, and 7.2.1 (2 outliers); 2.2.3, 6.3.3, 6.3.4, 7.3.2, and 7.3.4 (3 outliers); 6.3.1 (4 outliers); and 4.1.3, 5.2.5, 6.1.1, and 7.2.5 (5 outliers).

12 This distributional issue affects variables 2.3.3, 4.3.3, 5.3.4, 6.1.2, and 6.3.2 (factor *f* of 1).

13 These formulas achieve two things: converting all series into "goods" and scaling the series to the range [1, max] so that natural logs are positive starting at 0. Where "min" and "max" are the minimum and maximum indicator sample values.

The corresponding formula for bads is:

$$\ln\left[\frac{(Max \times f - 1)(Max - economy\ value)}{Max - Min} + 1\right]$$

References:

Groeneveld, R. A., & Meeden, G. (1984). Measuring Skewness and Kurtosis. *The Statistician, 33*, 391–99.

Paruolo P., Saisana, M., & Saltelli, A. (2013). Ratings and Rankings: Voodoo or Science? *Journal of the Royal Statistical Society, A 176*(2), doi: 0964–1998/13/176000.