

Appendix I The Global Innovation Index rational and origins, its conceptual framework and data limitations

Rationale and origins

The Global Innovation Index (GII) was launched in 2007 (see Box Annex 1). The goal was to find and determine metrics and methods that could capture a picture of innovation in society that is as complete as possible.

There were several motivations for setting this goal. First, innovation is important for driving economic progress and competitiveness – both for developed and developing economies. Many governments are putting innovation at the center of their growth strategies. Second, the definition of innovation has broadened – it is no longer restricted to research and development (R&D) laboratories and published scientific papers. Innovation is more general and horizontal in nature, and includes social, business model and technical aspects. Last, but not least, recognizing and celebrating innovation in emerging markets is critical for inspiring people – especially the next generation of entrepreneurs and innovators.

Box Annex 1: History of the GI (2007–2021)

The GI project was launched by Professor Soumitra Dutta in 2007 during his tenure at INSEAD. WIPO started its association with the GI in 2011 and began co-publishing the GI in 2012. In 2013, Cornell University joined as co-publisher, with Professor Dutta representing the GI at Cornell University and Bruno Lanvin at INSEAD. The GI continued to be co-published by Cornell University, INSEAD and WIPO up to 2020. As of 2021, the GI is published by WIPO in partnership with the Portulans Institute, various corporate and academic network partners and the GI Advisory Board.

Now in its 14th edition, the GI helps to create an environment in which innovation factors are under continual evaluation. It provides a key tool for decision-makers and a rich database of detailed metrics that are convenient for refining innovation policies.

Measuring innovation outputs and their impact remains difficult, hence great emphasis is placed on measuring the climate and infrastructure for innovation and on assessing related outcomes.

Although the final results take the shape of several rankings, the GI is more concerned with improving the “journey” to better measurement, understanding innovation, and identifying targeted policies, good practices and other levers that foster innovation. The rich data metrics, at index, sub-index or indicator level, can be used to monitor performance over time and to benchmark developments against economies within the same region or income group classification.

Defining innovation in the GI

The GI adopts a broad notion of innovation, originally elaborated in the *Oslo Manual* developed by the European Communities and the Organisation for Economic Co-operation and Development (OECD). In its fourth edition, the *Oslo Manual 2018* introduces a more general definition of innovation:¹

An innovation is a new or improved product or process (or combination thereof) that differs significantly from the unit's previous products or processes and that has been made available to potential users (product) or brought into use by the unit (process).

This update of the *Oslo Manual* also introduces a series of definitions associated with innovation in business activities and for different types of innovation firms. In this context, innovation translates as improvements made to outcomes in the form of either new goods or services or any combination of these. While the GI focuses on a more general definition of innovation, it is important to highlight how these definitions capture the evolution of the way innovation has been perceived and understood over the last two decades.

Economists and policymakers previously focused on R&D-based technological product innovation, largely produced in-house and mostly in manufacturing industries. Innovation of this nature was executed by a highly educated labor force in R&D-intensive companies. The process leading to such innovation was conceptualized as closed, internal and localized. Technological breakthroughs were necessarily “radical” and took place at the “global knowledge frontier.” This characterization implied the existence of leading and lagging economies, with low- or middle-income economies only playing “catch up.”

Today, innovation capability is increasingly seen as the ability to exploit new technological combinations; it embraces the notion of incremental innovation and “innovation without research.” Non-R&D innovative expenditure is an important component of reaping the rewards of technological innovation. Interest in understanding how innovation evolves in low- and middle-income economies is increasing, along with an awareness that incremental forms of innovation can impact development. Furthermore, the process of innovation itself has changed significantly. Investment in innovation-related activity and intangible assets has consistently intensified at the firm, economy and global levels, adding both new innovation actors from outside high-income economies and non-profit actors. The structure of knowledge production activity is more complex and geographically dispersed than ever.²

A key challenge is to find metrics that capture innovation as it actually happens in the world today. Direct official measures that quantify innovation outputs remain extremely scarce. For example, there are no official statistics on the amount of innovative activity – defined as the number of new products, processes, or other innovations – for any given innovation actor, let alone for any given country (see the GII 2013, Chapter 1, Annex 1, Box 1). Most measurements also struggle to appropriately capture the innovation outputs of a wider spectrum of innovation actors, such as the services sector or public entities. This includes innovation surveys, which have contributed greatly to the measurement of innovation activities, but fail to provide a good and reliable sense of cross-economy innovation output performance, and are often not applicable to developing economies where innovation is often informal.³

The GII aims to improve the measurement of innovation in order to provide a more complete picture of innovation ecosystems across the globe.

The GII conceptual framework

The overall GII ranking is based on two sub-indices that are both equally important in presenting a complete picture of innovation; the Innovation Input Sub-Index and the Innovation Output Sub-Index. Hence, three indices are calculated:

- **Innovation Input Sub-Index:** Five input pillars capture elements of the economy that enable and facilitate innovative activities.
- **Innovation Output Sub-Index:** Innovation outputs are the result of innovative activities within the economy. Although the Output Sub-Index includes only two pillars, it carries the same weight as the Input Sub-Index in calculating the overall GII scores.
- **The overall GII score** is the average of the Input and Output Sub-Indices, on which the GII economy rankings are then produced.

Each of the five input and two output pillars is divided into three sub-pillars, each of which is composed of individual indicators, a total of 81 this year (see the Economy profiles section for the Framework of the Global Innovation Index 2021). A deeper elaboration of the conceptual framework and pillars can be found in last year’s edition.⁴ Sub-pillars are calculated using the weighted average of its individual indicators and are normalized to take the form of scores between 0 and 100. Pillar scores are calculated using the weighted average of its sub-pillar scores.

Adjustments to the GII model in 2021

Annex Table 1 summarizes adjustments to the GII 2021 framework. A total of 11 indicators were modified this year. The methodology of five indicators changed, three are new indicators, two indicators were dropped, and one indicator changed name.

Annex Table 1
Changes to the GII 2021 framework

GII 2020	Adjustment	GII 2021
4.2.3 Venture capital deals/bn PPP\$ GDP	Methodology revised	4.2.3 Venture capital investors, deals/bn PPP\$ GDP
	New indicator	4.2.4 Venture capital recipients, deals/bn PPP\$ GDP
4.3.2 Intensity of local competition [†]	Removed	
	New indicator	4.3.2 Domestic industry diversification
5.2.4 JV–strategic alliance deals/bn PPP\$ GDP	Methodology revised	5.2.4 Joint venture/strategic alliance deals/bn PPP\$ GDP
6.1.4 Scientific & technical articles/bn PPP\$ GDP	Methodology revised	6.1.4 Scientific and technical articles/bn PPP\$ GDP
6.2.1 Growth rate of PPP\$ GDP/worker, %	Indicator name changed	6.2.1 Labor productivity growth, %
6.2.5 High- & medium-high-tech manufacturing, %	Methodology revised	6.2.5 High-tech manufacturing, %
	New indicator	6.3.2 Production and export complexity
6.3.2 High-tech net exports, % total trade	Methodology revised	6.3.3 High-tech exports, % total trade
6.3.4 FDI net outflows, % GDP	Removed	

Source: Global Innovation Index 2021, WIPO.

Notes: Refer to the Sources and definitions (Appendix III) for a detailed explanation of terminology and acronyms.

Data limitations and treatment

This year the GII model includes 132 economies, which represent 94.3% of the world’s population and 99.0% of the world’s GDP in purchasing power parity current international dollars.

The timeliest possible indicators are used for the GII 2021: from the non-missing data, 30.0% are from 2020, 41.4%

are from 2019, 17.5% are from 2018, 5.9% are from 2017, 1.2% are from 2016, and the small remainder of 4.0% are from earlier years.⁵

The GII 2021 model includes 81 indicators, which fall into three categories:

- quantitative/objective/hard data (63 indicators);
- composite indicators/index data (15 indicators); and
- survey/qualitative/subjective/soft data (3 indicators).

This year, for an economy to feature in the GII 2021, the minimum symmetric data coverage is at least 36 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 the GII 2021, 132 economies had sufficient data available to be included in the Index. For each economy, only the most recent yearly data were considered. As a rule, the GII indicators consider data from as far back as 2011, with a few exceptions.

Missing values

For the sake of transparency and replicability of results, missing values are not estimated; they are indicated with “n/a” and are not considered in the sub-pillar score. In return, the European Commission’s Competence Centre on Composite Indicators and Scoreboards at the Joint Research Centre (JRC-COIN) audit (see Appendix II) assesses the robustness of the GII modeling choices (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 II).

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. Only hard data indicators were treated (32 out of 63).

First rule: selection

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

- an absolute value of skewness greater than 2.25; and
- a kurtosis greater than 3.5.⁶

Second rule: treatment

Indicators with one to five outliers (30 cases) were winsorized; the values distorting the indicator distribution were assigned the next highest value, up to the level where skewness and/or kurtosis had the values specified above.⁷

Indicators with five or more outliers and for which skewness or kurtosis did not enter within the ranges specified above were transformed using natural logarithms after multiplication by a given factor f .⁸ Since only “goods” were affected (i.e., indicators for which higher values indicate better outcomes, as opposed to “bads”), the following formula was used:

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

where “min” and “max” are the minimum and maximum indicator sample values.⁹

Normalization

The 81 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 the minimum and maximum indicator sample values, respectively. Index and survey data were exceptions; the original series range of values was kept as min and max values ([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 formulas were applied:

$$\text{Goods: } \frac{economy\ value - min}{max - min} \times 100$$

$$\text{Bads: } \frac{max - economy\ value}{max - min} \times 100$$

Caveats on the year-to-year comparison of 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. Each ranking reflects the relative positioning of a particular economy based on the conceptual framework, the data coverage and the sample of economies of that GII edition, 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 (changes in indicator composition and measurement revisions);
- 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 rankings or scores:

- **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, in an attempt to limit the number of missing data points.
- **Normalization factor.** Most GII variables are normalized using either GDP or population, with the intention of enabling cross-economy comparability. Yet, this implies that year-on-year changes in individual indicators may be driven either by the variable (numerator) or by its normalization factor (denominator).
- **Consistent data collection.** Measuring the change in 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.

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

Notes:

- 1 Eurostat and OECD, 2018.
- 2 See WIPO (2011–2021) for bi-annual elaborations on the changing nature and geographic dispersion of innovation. See Arundel et al. (2021) for an elaboration on the role and measurement of knowledge and technology transfer between innovation actors.
- 3 On innovation in the informal economy, see Kraemer-Mbula and Wunsch-Vincent (2017).
- 4 See WIPO (2020), Appendix 1: https://www.wipo.int/edocs/pubdocs/en/wipo_pub_gii_2020-appendix1.pdf.
- 5 The GII is calculated based on 9,647 data points out of a possible 10,692 (132 economies multiplied by 81 indicators), implying that 9.8% of data points are missing. The Sources and Definitions (Appendix III) include the range of years used for each indicator. If an indicator for an economy is missing, it is marked as “n/a” in the Economy profiles.
- 6 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 (132 economies).
- 7 This distributional issue affects the following variables: 2.1.5, 3.2.1, 4.2.2, 5.2.3, 5.2.4, 5.3.2, 5.3.3, 5.3.4, 6.1.5, 7.2.2, 7.2.4 and 7.3.1 (1 outlier); 2.2.3, 5.3.1 and 7.1.3 (2 outliers); 4.2.4, 6.1.3, 6.3.4, 7.1.1, 7.2.1, 7.3.2 and 7.3.4 (3 outliers); 5.2.5, 6.3.1 and 7.2.5 (4 outliers); and 4.2.3, 6.1.1, 6.1.2 and 6.3.3 (5 outliers). An exception was made this year by also winsorizing an indicator that had six outliers: 4.1.3.
- 8 Indicators 2.3.3 and 4.3.3 were treated using log-transformation (factor f of 1).
- 9 This formula achieves 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:

$$\ln \left[\frac{(\max \times f - 1) \times (\max - \text{economy value})}{\max - \min + 1} \right]$$

References

Arundel, A., S. Athreye and S. Wunsch-Vincent (2021). *Harnessing Public Research for Innovation in the 21st Century: An International Assessment of Knowledge Transfer Policies*. Series: Intellectual Property, Innovation and Economic Development. Geneva and Cambridge: World Intellectual Property Organization and Cambridge University Press.

Cornell University, INSEAD and WIPO (World Intellectual Property Organization) (2013). *The Global Innovation Index 2013: The Local Dynamics of Innovation*, S. Dutta and B. Lanvin (eds). Ithaca, NY, Fontainebleau and Geneva: Cornell, INSEAD and WIPO.

Eurostat and OECD (Organisation for Economic Co-operation and Development) (2018). *Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data* (4th ed.). Paris: OECD. <https://doi.org/10.1787/9789264304604-en>.

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

Kraemer-Mbula E. and S. Wunsch-Vincent (2017). *The Informal Economy in Developing Nations: Hidden Engine of Innovation?* Series: Intellectual Property, Innovation and Economic Development. Geneva and Cambridge: WIPO and Cambridge University Press.

WIPO (World Intellectual Property Organization) (2011–2021). *World Intellectual Property Report*, various editions. Geneva: WIPO.