

Mobile Connectivity Index Methodology

June 2024

Copyright © 2024 GSMA

GSMA

GSMA

The GSMA is a global organisation unifying the mobile ecosystem to discover, develop and deliver innovation foundational to positive business environments and societal change. Our vision is to unlock the full power of connectivity so that people, industry and society thrive. Representing mobile operators and organisations across the mobile ecosystem and adjacent industries, the GSMA delivers for its members across three broad pillars: Connectivity for Good, Industry Services and Solutions, and Outreach. This activity includes advancing policy, tackling today's biggest societal challenges, underpinning the technology and interoperability that make mobile work, and providing the world's largest platform to convene the mobile ecosystem at the MWC and M360 series of events.

We invite you to find out more at gsma.com

Follow the GSMA on Twitter: @GSMA

GSMA Connected Society

The Connected Society programme works with the mobile industry, technology companies, the development community and governments to increase access to and adoption of mobile internet, focusing on underserved population groups in developing markets.

For more information, please visit:
www.gsma.com/connected-society

GSMA Intelligence

GSMA Intelligence is the definitive source of global mobile operator data, analysis and forecasts, and publisher of authoritative industry reports and research. Our data covers every operator group, network and MVNO in every country worldwide — from Afghanistan to Zimbabwe. It is the most accurate and complete set of industry metrics available, comprising tens of millions of individual data points, updated daily.

GSMA Intelligence is relied on by leading operators, vendors, regulators, financial institutions and third-party industry players, to support strategic decision-making and long-term investment planning. The data is used as an industry reference point and is frequently cited by the media and by the industry itself.

Our team of analysts and experts produce regular thought-leading research reports across a range of industry topics.

www.gsmaintelligence.com
info@gsmaintelligence.com

Published: June 2024

Contents

1. Introduction **4**
About the Mobile Connectivity Index 4
Theoretical framework 4

2. Data selection **6**
Selection criteria 6
Indicators 7
Indicator updates 10
Pricing data 11
Taxation 13
Language accessibility of top ranked apps 13

3. Data treatment **14**

4. Normalisation **17**

5. Weighting **18**

6. Aggregation **22**

7. Uncertainty analysis **23**

8. Feedback **26**

1. Introduction

About the Mobile Connectivity Index

In 2022, more than 200 million people connected to mobile internet for the first time, bringing the total connected population to more than 4.5 billion people globally (or 57% of the global population).[1] For many of these individuals, mobile is the only method of accessing the internet, so growth in mobile internet adoption also drives digital inclusion, especially in low- and middle-income countries (LMICs).[2]

Mobile internet connectivity is not, however, equitable; in 2022, almost 50% of the LMIC population (more than 3 billion people) were connected, compared to 85% of the population in high-income countries, while in least developed countries (LDCs) only 25% were connected. Almost half the world's population are therefore still unable to realise the social and economic benefits that mobile internet can enable.

To support the mobile industry's efforts to drive mobile internet connectivity and accelerate digital inclusion, the Mobile Connectivity Index (MCI) measures and tracks enablers of mobile internet connectivity. The MCI has been constructed according to the steps set out in the guidelines developed by the OECD and the European Commission's Joint Research Centre (JRC).[3] This methodology presents the theoretical framework that underpins the MCI; the process for selecting the indicators, along with how they are structured; the approach used to normalise the data; the weights used in the MCI; and the approach to aggregation.

Theoretical framework

What is measured?

The MCI measures the enablers of mobile internet connectivity. It is therefore an input index. An input index measures a number of indicators that lead to an important outcome, in this case why people are (or are not) using mobile internet connectivity. This is different to an output index, which, in the context of mobile connectivity, might seek to measure and understand how (or how much) people are using mobile internet services.[4]

Why is an index necessary?

There is no single barrier or enabler to mobile internet connectivity; rather, a number of prerequisites are necessary for a country's population to use mobile internet services. An index is required because it measures multiple enablers and can summarise the complex reality of delivering mobile internet connectivity in a country.

1. Source: GSMA Intelligence
2. This categorisation is based on the World Bank's country classification and includes countries that the World Bank classifies as low-income, lower-middle income and upper-middle income.
3. Handbook on constructing composite indicators: methodology and user guide, OECD and JRC, 2008
4. In order to track outcomes and the impact of mobile technology, the GSMA publishes the annual Mobile Industry SDG Impact report.

A number of indices exist in the ICT sector[5] and so the MCI has been designed to ensure that it does not replicate other related indices. In this respect, the MCI has four key characteristics that together distinguish it from other indices:

- It focuses specifically on mobile internet connectivity rather than internet connectivity in general (including fixed-line connectivity). Given that the internet gap in LMICs is expected to be addressed to a significant extent by mobile, it is important to understand the enablers of mobile connectivity specifically.
- As the MCI is focused on mobile internet connectivity, the majority of the underlying indicators are unique to the MCI and are either not available or not used in other indices (e.g. spectrum availability, taxation on mobile services and the availability of content on mobile platforms).
- It is an input index that seeks to measure the performance of countries against a set of key enabling characteristics, rather than an output index that measures internet take-up and usage.
- It is a global index, encompassing 170 countries that account for more than 99% of the world's population.

The MCI scores are highly correlated with mobile internet adoption, and countries that achieve significant improvements on the MCI are also more likely to have seen increases in mobile internet adoption over time.[6] This means that the MCI is an effective tool to identify priorities to drive mobile internet adoption.

How are the enablers measured?

The enablers of mobile internet connectivity that inform the indicators selected for the MCI are:

1. **Infrastructure:** The availability of high-performance mobile internet network coverage.
2. **Affordability:** The availability of mobile services and devices at price points that reflect the level of income across a national population.
3. **Consumer Readiness:** Citizens with the awareness and skills needed to value and use the internet.
4. **Content and Services:** The availability of secure online content and services accessible and relevant to the local population.

Changes to the 2023 Mobile Connectivity Index

As part of the update to the MCI in 2023, some methodological changes were made, and these have been retained in the 2024 MCI. These were based on feedback we have received, as well as the statistical audit that was carried out by the European Commission's JRC.[7] Furthermore, changes in the availability of data necessitated updates to some indicators. Updates to the methodology applied since 2023 are covered in the relevant section of this document. In order to ensure consistency and comparability, all changes are applied to each year (from 2014 to 2023). This means that changes in scores over time can be compared on a like-for-like basis in the 2024 MCI. However, there will be differences in the 2024 version compared with previous MCI publications.

5. For example, the Digital Economy and Society Index (European Commission), the Inclusive Internet Index (Meta and the Economist Intelligence Unit) and the Network Readiness Index (Portulans Institute).
6. See State of Mobile Internet Connectivity 2018, GSMA
7. See JRC statistical audit of the Mobile Connectivity Index

2. Data selection

Selection criteria

As the MCI is an input index, it is important that each indicator is an input for mobile internet connectivity rather than an output or outcome (e.g. measuring the level of take-up). It is also important to develop a set of criteria against which each indicator can be considered for inclusion in the MCI. The following criteria have therefore been developed, based on guidance from the JRC and OECD:

- **Relevance:** The indicator should measure a barrier or an enabler in the take-up of mobile internet services.
- **Accuracy:** The indicator should correctly estimate or describe the quantities or characteristics they are designed to measure.
- **Coverage:** The data should cover as many countries as possible, as the MCI is intended to be a global index. An indicator is not included if there is missing data on more than 25% of countries in the MCI.
- **Timeliness:** The data should be collected consistently over time.

A key consideration in the assessment of accuracy is to include, to the greatest extent possible, 'hard' indicators that are objective and can be quantified. These are distinct from 'soft' indicators that are usually based on qualitative data from surveys or case studies; such indicators are typically used to measure things that are difficult to quantify – for example, the quality of governance or the extent of corruption. Although soft indicators are very useful for some indices, particularly those where hard indicators are difficult to develop, they are not used in the MCI. This is to ensure that operators, regulators and other stakeholders have objective benchmarks on which to target improved performance.

Although the indicators included in the MCI have all been carefully chosen based on the above criteria, there are some cases where data constraints require the use of proxy indicators if it is not possible to measure an enabler with complete accuracy:

- There is currently no data comparing a large number of countries in the area of digital skills or awareness, particularly on a mobile platform. More traditional skills indicators (e.g. literacy and expected years of schooling) are therefore used to measure consumers' ability to effectively use and engage with digital technology.
- Indicators measuring the availability of mobile-specific content primarily focus on smartphone applications due to the lack of comparable data across countries on other types of mobile content.

Although the vast majority of the indicators are highly correlated with mobile internet penetration, suggesting that on average they are associated with higher take-up, there may be specific countries where they work less well as proxy indicators. These indicators will therefore be assessed going forward and, where they can be improved, incorporated into future versions of the MCI.

Indicators

Table 1 details the indicators that make up the MCI. The MCI comprises four key enablers, which in turn comprise a number of dimensions. These dimensions are constructed by aggregating one or more indicators.

Table 1
Mobile Connectivity Index enablers, dimensions and indicators

Enabler	Dimension	Indicator	Original unit of measurement	Source
Infrastructure	Network coverage	2G population coverage	% of population covered	ITU World Indicators/ICT Indicators database
		3G population coverage	% of population covered	GSMA Intelligence
		4G population coverage	% of population covered	GSMA Intelligence
		5G population coverage[a]	% of population covered	GSMA Intelligence
	Network performance	Mobile download speeds	Mbps	Ookla® Speedtest Intelligence®
		Mobile upload speeds	Mbps	Ookla Speedtest Intelligence
		Mobile latencies	Milliseconds	Ookla Speedtest Intelligence
	Spectrum	Spectrum assigned in bands below 1 GHz	MHz assigned to mobile operators	GSMA Intelligence
		Spectrum assigned in bands between 1–3 GHz	MHz assigned to mobile operators	GSMA Intelligence
		Spectrum assigned in bands between 3–6 GHz[a]	MHz assigned to mobile operators	GSMA Intelligence
		Spectrum assigned in mmWave bands[a]	MHz assigned to mobile operators	GSMA Intelligence

Table 1 continued

Mobile Connectivity Index enablers, dimensions and indicators

Enabler	Dimension	Indicator	Original unit of measurement	Source
Affordability	Mobile data affordability	Affordability of entry basket	% of average monthly GDP per capita	Tarifica
		Affordability of higher basket	% of average monthly GDP per capita	Tarifica
		Affordability of entry basket for poorest 40%	% of monthly GDP per capita for poorest 40%	Tarifica
		Affordability of higher basket for poorest 40%	% of monthly GDP per capita for poorest 40%	Tarifica
	Handset affordability	Device affordability	% of average monthly GDP per capita	Tarifica
		Device affordability for poorest 40%	% of monthly GDP per capita for poorest 40%	Tarifica
	Taxation	Cost of taxes on mobile data	Cost of tax as a % monthly data price	GSMA Intelligence
		Cost of taxes on handsets	Cost of tax as a % of handset price	GSMA Intelligence
		Cost of sector-specific taxes on mobile data	Cost of mobile-specific taxes as a % monthly data price	GSMA Intelligence
Consumer Readiness	Basic skills	Adult literacy rate	% of literate adult population (above 15 years old)	UN and UNESCO
		School life expectancy[(b)]	Years	UN
	Gender equality	Gender gap in mobile ownership	1 – female adoption/male adoption	GSMA Intelligence and Gallup World Poll
		Gender gap in mobile internet use	1 – female adoption/male adoption	GSMA Intelligence, Gallup World Poll and DataReportal
	Mobile ownership	Penetration of mobile users	Unique mobile subscribers as a % of population	GSMA Intelligence

Table 1 continued

Mobile Connectivity Index enablers, dimensions and indicators

Enabler	Dimension	Indicator	Original unit of measurement	Source
Content and Services	Local relevance	Top-level domains (TLDs) per person[(c)]	Number of domains per 100 people	ZookNIC
		E-Government score[(d)]	Index value (0–1)	UN
		Mobile social media penetration[(e)]	Mobile social media accounts as a % of population	DataReportal
		Locally developed apps per person	Number of active mobile apps developed per person	Appfigures
		Digital language support	Digital language support score	Derivation
		Language accessibility of top ranked apps	Average of the % of population that can use each app in the top 400 for the country	Appfigures, Ethnologue and GSMA Intelligence
	Online security	Global Cybersecurity Index	Index value (0–1)	ITU

(a) The indicators for 5G coverage and the assignment of spectrum above 3 GHz are incorporated from 2021, as 5G technology was not deployed before 2019 and was only rolled out in a limited number of countries in 2019–2020.

(b) This is the total number of years of schooling (primary to tertiary) that a child can expect to receive given current enrolment rates. It is therefore a forward-looking indicator.

(c) This includes the number of generic top-level domains (gTLDs) registered in a country and the number of registered country-code top-level domains (ccTLDs).

(d) This indicator uses the Online Service Index score in the UN's E-Government Survey.

(e) Due to revisions to data published by various social media platforms, there are some years when data on mobile social media penetration significantly changes due to differences in reporting rather than actual social media use.[8]

8. For further details, see https://datareportal.com/notes-on-data

Indicator updates

As part of the methodology update in 2023, several indicators were removed from the MCI that were included in previous editions:

- The dimension 'Other enabling infrastructure' was removed, along with each of the four indicators (international bandwidth per user, number of secure servers, access to electricity and number of internet exchange points). This dimension was designed to capture the quality of a country's backhaul and core network, which are needed to deliver a high-quality end-to-end mobile service. However, some indicators are no longer updated regularly and so do not capture the characteristics needed. We considered more relevant indicators such as the proportion of base stations connected with fibre, and the amount of capacity built into mobile networks. Unfortunately, there is insufficient comparable data on these metrics across a large number of countries. If this changes going forward, we will incorporate new metrics where they enhance the MCI.

- The dimension and indicator 'Inequality' was removed. High levels of income inequality can have important implications for the adoption of mobile internet services. For example, if two countries have an average (mean) income of $1,000 per month but one has perfect income equality and the other has high levels of inequality, then the cost of a device or mobile data will be much less affordable for a large proportion of the population in the second country. We now capture this enabler more directly by including affordability measures of mobile data and devices for the poorest 40% of the population in each country.

- Two skills indicators on mean years of schooling and tertiary enrolment have been removed, as they were highly correlated with expected years of schooling. Furthermore, given that these metrics were proxies for digital skills (rather than direct measures), the removal of two indicators simplifies and streamlines the MCI.

- Four indicators on gender equality have been removed: gender parity ratios for mean years of schooling, GNI per capita, financial account access and the Women, Business and the Law index. These were originally included in the MCI because they can be used to understand some of the reasons for a high mobile gender gap in a given country. However, the two most important and relevant indicators in this dimension are those related to mobile internet adoption, namely the gender gaps in mobile ownership and mobile internet use. As part of our attempts to streamline the MCI, we have therefore focused this dimension on the indicators specific to mobile internet.

- The indicator on the number of apps in the national language has been removed, as it was highly correlated with the language accessibility of top ranked apps. We have also integrated the 'Availability' dimension into the 'Local relevance' dimension within the Content and Services enabler.

We also made adjustments to the following indicators:

- Spectrum indicators are now based on the total amount of spectrum assigned to operators in each country, rather than the average spectrum per operator.

- Mobile data affordability is now measured based on two baskets rather than four baskets. In 2014–2018, the 'entry' basket is based on a consumer using 100 MB of data per month and the 'higher' basket is based on 500 MB of data. In 2019–2023, the entry basket is based on 1 GB of data and the higher basket is based on 5 GB. This approach is more flexible and reflects the increasing use of data by consumers over time,[9] rather than continuing to track baskets that no longer reflect current consumption patterns.

- The taxation indicators now consider mobile data and handsets separately, rather than combining them into a 'total cost of mobile ownership' measure.

Lastly, a new indicator has been added to the Content and Services enabler on digital language support. This metric is sourced from Derivation[10] and provides a measure of the aggregated digital capabilities for all living languages within a country, based upon the availability and accessibility of language-specific hardware/software support.[11]

9. For further analysis on this, see the State of Mobile Internet Connectivity Report 2022.
10. https://derivation.co/
11. Further details on this indicator can be found in Simons et al (2022), Assessing Digital Language Support on a Global Scale

Pricing data

Pricing data is provided by Tarifica, with retail prices captured as of the first quarter of 2017, 2018, 2019, 2020, 2021, 2022 and 2023, including all relevant taxes.

Mobile data

In order to produce comparable price metrics across countries, four baskets are defined based on usage allowance, contract and technology. The baskets were designed to capture entry usage as well as higher usage for more intense users. In order to construct the baskets, the following information was taken into account:

- Historical trends in average data consumption across countries, sourced from GSMA Intelligence, Ofcom, Tefficient and Opera. Future data requirements (which are likely to increase) were taken into account. We also gave due consideration to the fact that average values are often distorted by particularly intensive users of mobile services. This means that median data usage is likely to be lower than average data usage.[12]
- A selection of allowances currently offered by operators in developed and emerging markets, provided by Tarifica.
- Baskets used in other mobile pricing benchmark studies published by the OECD, Tarifica, Ofcom, the European Commission and the ITU. These represent basket designs often used in economics literature that analyse pricing in the mobile industry.

The baskets resulting from this analysis are summarised in Table 2.

Table 2
Usage basket profiles

Basket	Entry (2014–2018)	High (2014–2018)	Entry (2019–2023)	High (2019–2023)
Monthly usage allowance	100 MB data	500 MB data	1 GB data	5 GB data
Tariff	Prepaid	Any	Any	Any
Technology	2G, 3G or 4G	3G or 4G	3G or 4G	3G or 4G

Having defined these baskets, Tarifica researched all tariffs offered by operators in each country and selected the cheapest available plan under which the basket requirements could be met. In addition, the following guidelines were applied to ensure prices were representative of regular usage and consumption patterns:

- Prepaid plans lasting less than one month were included – in such cases, the usage allowance and price were scaled up to one month to ensure comparability across tariffs (e.g. the usage and price of a five-day plan were multiplied by six to derive a monthly usage and price).
- Short-term promotional offers were excluded.
- Plans targeted at or restricted to certain profiles (e.g. youth, student, senior) were not included.
- Where a tariff included an initial one-off fee (e.g. activation, SIM card), this was amortised over the contract length or validity period.

12. See the State of Mobile Internet Connectivity Report 2022.

Prices have been captured in local currencies for all countries for the years 2017–2023, before converting to US dollars. To derive pricing data for 2014–2016, we incorporated data from the ITU, which has historically (before 2018) collected data on the 500 MB basket.[13] Using ITU pricing data for 2014–2017, we calculated growth rates for 2016–2017, 2015–2016 and 2014–2015. These growth rates were applied to the 2017 prices obtained from Tarifica to give a complete historical dataset for each basket. Lastly, for some countries in 2017–2022 that were not included in the Tarifica dataset, we incorporated historical pricing data published by the Alliance for Affordable Internet.[14]

Handset price

As the MCI is particularly focused on connecting the unconnected, we are interested in measuring prices of entry-level handsets that allow users access to the internet rather than high-end devices that are sold at premium prices. In order to obtain this data, Tarifica researched the cheapest handset available in each market with internet-browsing capability. This device could therefore be a smartphone,[15] smart feature phone[16] or feature phone.[17] Given that the performance for basic internet mobile applications (such as basic video or social networking) is only functional with 3G and 4G, this analysis excluded devices with 2G and WAP connectivity. Device prices were collected by researching the devices available on the websites of all MNOs in each country; other retailers' websites were analysed for the countries where MNOs did not offer handsets. This approach was taken due to the significant resource that would be required in order to inspect all non-MNO retailers in each country. This means that in some markets there may be cheaper devices available, but by keeping the approach consistent across countries, the relative differences across countries should be similar and minimise the risk of bias.

As with mobile tariffs, prices were captured in local currencies before converting to US dollars for the years 2017–2023. To derive pricing data for 2014–2016, we used data from Strategy Analytics on the average selling price (ASP) of handsets for each country and estimated historical prices by applying ASP growth rates to 2017 data. Similar to the mobile pricing data, for some countries in 2017–2022 that were not included in the Tarifica dataset, we incorporated device pricing data published by the Alliance for Affordable Internet.[18]

13. See the ITU Price Baskets. The ITU's approach to collecting pricing data differs to the approach used by Tarifica, particularly because the ITU only collects data from the largest mobile operator whereas Tarifica considers all operators.
14. See https://a4ai.org/research/mobile-broadband-pricing/
15. A smartphone is a device that enables advanced access to internet-based services and other digital functions. Smartphone platforms, such as Android, iOS, Windows Phone and BlackBerry, support a broad range of applications created by third-party developers.
16. A smart feature phone has an operating system that supports a range of applications created by third-party developers and that are formatted to work on a smaller screen and accessed via a 9-key layout, not a touch screen.
17. A feature phone allows basic access to internet-based services but on a closed platform that does not support a broad range of applications. The handset supports additional features such as a camera and the ability to play multimedia files such as music and video.
18. Alliance for Affordable Internet (2022). 2022 prices and affordability of smartphones and feature phones by country [database]. Retrieved from https://a4ai.org/research/device-pricing-2022/

Taxation

The taxation indicators are developed by estimating the proportion of the monthly cost of mobile data that are (i) accounted for by all taxes and (ii) accounted for by mobile-specific taxes. The third indicator is the proportion of the total cost of a device that is accounted for by all taxes. The reference prices used for mobile data are the 500 MB basket from 2014 to 2018 and the 1 GB basket from 2019 to 2023. The reference price for the device is the same price gathered as part of the indicators on handset affordability.

In the case of mobile data, the following taxes are included:

- General sales or ad valorem taxes (e.g. VAT)
- Sector-specific ad valorem taxes (e.g. mobile excise duties)
- Sector-specific monthly charges on the use of mobile data

In the case of handsets, the following taxes are included:

- General sales or ad valorem taxes (e.g. VAT)
- Handset-specific ad valorem taxes
- Customs duties (both ad valorem and fixed charges)
- Sector-specific charges to purchase a handset
- The activation and connection price or any other charges incurred to connect to the MNO's network. For prepaid customers, this usually consists of an initial charge for activating the SIM card. For postpaid customers, there may be additional upfront costs, such as an initial charge for activating the number.

Language accessibility of top ranked apps

Many people consume content that is produced outside of their own country, so it is important to measure the extent to which this more widely available content is accessible and relevant to users. We therefore developed an indicator to assess whether a country's population has content that they can understand and with which they can engage. We did this by using language as a measure for accessibility and relevance – if an individual has content available in a language they speak then they are more likely to find the internet useful.

The indicator requires data on the languages spoken in each country, which is sourced from Ethnologue. Specifically, a dataset was provided that had a list of languages spoken in each country as well as data on the proportion of population speaking each language, both as a first language and in total.

The second data source is a list of all mobile applications available on Google Play, Apple Store, Windows, Amazon and other smaller platforms. This dataset was provided by Appfigures. For each application, information is provided on the languages and countries it is available in as well as the application category (e.g. gaming, education, health), the year of release and whether or not it is active. Data was also provided on the top 400 ranked applications (based on downloads) for Google Play and Apple Store at the end of each year since 2014, giving an indication of the content most popular in each country.

The indicator on language accessibility of top ranked apps focuses on the most popular applications available in a country and measures the proportion of the population that can use them, whether in their first or second languages. For each application, we estimate the proportion of the population that are able to use it based on the languages that it features. If an application is available in English, French and Hindi and 80% of a country's population speaks one of these languages (either as a first or second language), we assume that the app is accessible to 80% of the population. We then take the average of the top 400 ranked applications in each store and use the most popular store for inclusion in the MCI (for example, Android is far more widely used in many markets than iOS, so for these countries we only consider the accessibility of mobile applications on Google Play).

3. Data treatment

Having collected data and carried out the necessary calculations for the above indicators, we check to ensure that each country has data on at least 75% of indicators overall and at least half the indicators within each enabler. This ensures that a significant proportion of data for each country is based on actual data rather than estimates. We also ensure that each indicator has data for at least 75% of countries.

The next step is to then 'treat' the data by dealing with outliers and imputing missing data. If data is skewed by certain outliers, this could impact the overall index scores (for example, a country with exceptionally low download speeds compared to all other countries will score very low but will also cause all other countries to score relatively highly with little variation). In order to identify outliers, indicators are assessed to see if they have an absolute skewness above 2 and kurtosis above 3.5.[19] Where these thresholds are met, one of two treatment approaches is adopted:

— **Winsorisation:** Outlier variables are trimmed to the nearest value until the indicator is brought within the specified ranges for skewness and kurtosis. For example, if a country has an outlier value of 1,000 and the next highest value is 90, the former is trimmed to 90. If this gives acceptable skewness and kurtosis scores, the process stops there. If not, the two values are trimmed to the next highest value (which might be 80 in the above example). This process is continued until the indicator falls within the specified skewness and kurtosis ranges. In order to ensure that a large number of observations are not adjusted, a maximum of six observations are trimmed. If this is still insufficient to reduce skewness and kurtosis, the second approach is implemented.

— **Transformation:** As the majority of the indicators with high skewness and kurtosis are skewed to the right, a logarithmic transformation is used to bring the indicator within the specified ranges.

There are some indicators where a logarithmic transformation is applied even though winsorisation would suffice. This is because a logarithmic transformation has a conceptual benefit in that it results in improvements in the lower end of the indicator distribution being more 'beneficial' to a country than improvements at the high end of the distribution. An example of this is in relation to mobile application development. Increasing the number of applications per 1,000 people from 10 to 20 is likely to have a bigger impact on providing relevant content than increasing the number from 100,010 to 100,020 and so – from the perspective of the MCI – should be rewarded with a higher increase. Logarithmic transformation achieves this.

19. These thresholds are generally used in identifying outliers for composite indices.

A logarithmic transformation has been applied to the following indicators, for either data treatment or conceptual reasons:

— All of the affordability metrics for mobile data and devices

— TLDs per person

— Locally developed apps per person.

The next step in the data treatment process requires the imputation of missing data. Where data is missing, historical information is used before implementing a modelling-based approach. If data is missing but there are values in preceding and subsequent years, then we apply linear interpolation. Otherwise, we extrapolate data as a constant value of the nearest reported data. This process is likely to result in a more accurate estimate for a specific country than using a modelled or imputed value based on data for other countries.

The remaining missing data is imputed with an expectation-maximisation with bootstrapping (EMB) multiple imputation algorithm. The EMB method generates imputed values through an iterative procedure that uses other variables to impute a value (expectation) and then asserts whether the value is most likely to fit the data (maximisation). To account for variation caused by missing data, the model is run 10 times – the average of these 10 imputations is then used to impute the missing value.[20]

In order to ensure that the MCI's rankings are robust to the imputation method, previous sensitivity analysis of the MCI also imputed missing values using a multiple imputation method based on predictive mean matching (PMM)[21] and based on a multivariate normal data augmentation procedure (MVN).[22] We found that the three imputation methods returned similar values and none has a specific pattern of overestimation or underestimation compared to the other two methods. Moreover, we observed that no countries moved more than eight places in the rankings and no countries had a variation of the MCI greater than five points when using any of the three approaches. This shows that the MCI is not particularly sensitive to the imputation methodology used for missing data.

20. The EMB imputation is done using the 'Amelia II' package in R.
21. This generates estimates of missing values using a regression model; the independent explanatory variables are selected if they have a high correlation with the variable being imputed. The model generates an estimated value using the regression for a country that is missing data and then matches it with the three nearest countries with the closest regression output. The average values across those countries is then taken.
22. This generates estimates of missing values assuming the variables used to impute and the imputed variable jointly follow a multivariate normal distribution.

Table 3

Number and percentage of multiple imputed data points by indicator

Variable	# imputed	% imputed
2G Population Coverage	0	0.0%
3G Population Coverage	0	0.0%
4G Population Coverage	0	0.0%
5G Population Coverage	0	0.0%
Mobile download speeds	0	0.0%
Mobile upload speeds	0	0.0%
Mobile latencies	0	0.0%
Spectrum assigned in bands below 1GHz	0	0.0%
Spectrum assigned in bands between 1-3GHz	0	0.0%
Spectrum assigned in bands between 3-6GHz	0	0.0%
Spectrum assigned in mmWave bands	0	0.0%
Affordability of entry basket (1GB)	105	6.1%
Affordability of higher basket (5GB)	2	0.1%
Affordability of entry basket (1GB) for poorest 40%	106	6.1%
Affordability of higher basket (5GB) for poorest 40%	3	0.2%
Device affordability	37	2.1%
Device affordability for poorest 40%	38	2.2%
Cost of taxes on mobile data	0	0.0%
Cost of taxes on handsets	0	0.0%
Cost of sector specific taxes on mobile data	0	0.0%
Mobile ownership	0	0.0%
Literacy	0	0.0%
School Life Expectancy	0	0.0%
Gender gap in mobile ownership	230	13.3%
Gender gap in mobile internet	0	0.0%
Top-Level Domains (TLDs) per person	0	0.0%
E-Government Score	0	0.0%
Mobile Social Media Penetration	0	0.0%
Locally developed apps per person	0	0.0%
Digital Language Support	0	0.0%
Language accessibility of top ranked apps	170	9.8%
Global Cybersecurity Index	0	0.0%

4. Normalisation

Normalisation is required in an index to adjust for different units of measurement and different ranges of variation across the indicators. For the MCI, the minimum-maximum method is used, which transforms all indicators so that they lie within a range between 0 and 100 using the following formula:

$$I_{q,c} = \frac{X_{q,c} - min_c(X_q)}{max_c(X_q) - min_c(X_q)}$$

Where 'I' is the normalised min-max value, 'x' represents the actual value and the subscripts 'q' and 'c' represent the indicator and country respectively.

This method has been chosen over alternatives such as rankings and categorical scales because it retains interval-level information. For example, in the case of ranking 3G coverage, Country A might have 100% coverage, Country B might have 99% coverage and Country C might have 90% coverage. These countries would therefore be ranked in order as 1, 2 and 3 respectively (or they may all be categorised as having the highest score on an ordinal scale). However, this ranking does not take into account the differences between countries – specifically the fact that B is much closer to A than it is to C. Furthermore, as the MCI is updated over time, a ranking-based approach may not track a country's progress as well as minimum-maximum or standardisation because a country might improve its coverage without increasing its rank.

For most indicators, the minimum and maximum used for normalisation are based on the actual minimum and maximum for that indicator, although in some cases they have been amended. For example, the gender gap indicators have a minimum threshold of zero, as this represents gender equality. Any country with a value less than zero is therefore not rewarded with a higher score.

To allow for comparisons of index scores over time, the minimum and maximum for each indicator are fixed. Some of the indicator maxima have therefore been adjusted where there are likely to be increases during the next few years in order to give all countries room to improve. These adjustments are based on an analysis of historical data and statistical analysis (for example, ensuring that the maxima do not significantly exceed a threshold of being two standard deviations above the mean).

As part of the normalisation process, all indicators are transformed such that they have the same orientation i.e. a higher score always represents a 'better' score. This is necessary for indicators that are negatively correlated with mobile internet penetration, such as affordability, taxation and latency.

Previous sensitivity analyses of the MCI also showed that the rankings were not sensitive to other types of normalisation, for example z-scores.[23]

23. This transforms all indicators such that they have a mean of 0 and a standard deviation of 1.

5. Weighting

In order to aggregate indicators into dimension scores (and dimensions to enabler scores and enablers to an overall index score), it is necessary to assign a weight to each component of the MCI. To construct the weights at the dimension, enabler and overall index level, a number of considerations have been taken into account, including:

- Statistical relationship between indicators and dimensions concerning mobile internet penetration
- Analysis of consumer survey responses regarding perceived barriers to mobile internet access
- Principal component analysis to identify weights that correct for the overlapping information implied by grouping indicators that are correlated (rather than representing a measure of importance).
- Research carried out by the GSMA and other organisations on digital inclusion and barriers to mobile connectivity.
- Qualitative research.

Based on the above, the following weights have been used for the dimensions in the years 2021–2023 (Table 4), in the years 2014–2020 (Table 5) and enablers (Table 6).

Table 4

Indicator weights for dimensions for 2021–2023

Dimension	Indicator	Weight
Network coverage	2G population coverage	10%
	3G population coverage	40%
	4G population coverage	40%
	5G population coverage	10%
Network performance	Mobile download speeds	33.33%
	Mobile upload speeds	33.33%
	Mobile latencies	33.33%
Spectrum	Spectrum assigned in bands below 1 GHz	50%
	Spectrum assigned in bands between 1–3 GHz	30%
	Spectrum assigned in bands between 3–6 GHz	15%
	Spectrum assigned in mmWave bands	5%
Mobile data affordability	Affordability of entry basket	25%
	Affordability of higher basket	25%
	Affordability of entry basket for poorest 40%	25%
	Affordability of higher basket for poorest 40%	25%
Handset affordability	Device affordability	50%
	Device affordability for poorest 40%	50%
Taxation	Cost of taxes on mobile data	33.33%
	Cost of taxes on handsets	33.33%
	Cost of sector-specific taxes on mobile data	33.33%
Basic skills	Adult literacy rate	50%
	School life expectancy	50%
Gender equality	Gender gap in mobile ownership	50%
	Gender gap in mobile internet	50%
Mobile ownership	Penetration of mobile users	100%
Local relevance	Top-level domains (TLDs) per person	20%
	E-Government score	10%
	Mobile social media penetration	20%
	Locally developed apps per person	20%
	Digital language support	20%
	Language accessibility of top ranked apps	10%
Online security	Global Cybersecurity Index score	100%

Table 5

Indicator weights for dimensions for 2014–2020

Dimension	Indicator	Weight
Network coverage	2G population coverage	20%
	3G population coverage	40%
	4G population coverage	40%
Network performance	Network performance	33.33%
	Mobile upload speeds	33.33%
	Mobile latencies	33.33%
Spectrum	Spectrum	60%
	Spectrum assigned in bands between 1–3 GHz	40%
Mobile data affordability	Mobile data affordability	25%
	Affordability of higher basket	25%
	Affordability of entry basket for poorest 40%	25%
	Affordability of higher basket for poorest 40%	25%
Handset affordability	Handset affordability	50%
	Device affordability for poorest 40%	50%
Taxation	Taxation	33.33%
	Cost of taxes on handsets	33.33%
	Cost of sector-specific taxes on mobile data	33.33%
Basic skills	Adult literacy rate	50%
	School life expectancy	50%
Gender equality	Gender equality	50%
	Gender gap in mobile internet	50%
Mobile ownership	Mobile ownership	100%
Local relevance	Top-Level domains (TLDs) per person	20%
	E-Government score	10%
	Mobile social media penetration	20%
	Locally developed apps per person	20%
	Digital language support	20%
	Language accessibility of top ranked apps	10%
Online security	Global Cybersecurity Index score	100%

Table 6
Dimension weights for enablers

Enabler	Dimension	Weight
Infrastructure	Network coverage	40%
	Network performance	40%
	Spectrum	20%
Affordability	Mobile tariffs	40%
	Handset price	40%
	Taxation	20%
Consumer Readiness	Basic skills	33.3%
	Gender equality	33.3%
	Mobile ownership	33.3%
Content and Services	Local relevance	80%
	Online security	20%

In terms of weighting the enablers for the MCI, equal weights are assigned i.e. each enabler is given a weight of 25%. Table 7 shows the Pearson and Spearman ranking correlation coefficients between the enablers and overall index score against mobile internet penetration, demonstrating a high correlation across all enablers.

Table 7
Correlation coefficients with mobile internet penetration

Enabler	Pearson correlation	Spearman correlation
Infrastructure	0.86	0.87
Affordability	0.81	0.82
Consumer Readiness	0.87	0.91
Content and Services	0.92	0.92
Index	0.93	0.94

Analysis was conducted to assess the impact of adjusting these weights on the correlation between the overall index score and mobile internet penetration rates, including analysis that set weights to optimise both correlation coefficients. Such changes make very small improvements to the final index-penetration correlation (less than 0.01). Equal weights are therefore appropriate.

6. Aggregation

Once weights have been assigned to the indicators, dimensions and enablers, they need to be aggregated to produce the relevant composite scores. Two methods of aggregation were considered: arithmetic and geometric. The key consideration when choosing between these is the extent to which indicators, dimensions and enablers are substitutable, with arithmetic aggregation implying perfect substitutability and geometric implying partial substitutability.

When considering the indicators in the MCI, there is often a greater degree of substitutability compared to the dimensions and enablers. For example, within the network coverage dimension low 3G network coverage in a country can be compensated by high 4G network coverage. In the mobile data affordability dimension, an expensive price for the higher basket could be compensated by a cheap entry basket price. At the enabler level, such substitutability is unlikely to be perfect – a country with a high infrastructure score is unlikely to achieve high mobile internet penetration if mobile products or services are completely unaffordable or if there is no relevant internet content. The dimensions sit somewhere in between – there is likely to be more substitutability than the enabler level (e.g. high handset price might be compensated by low mobile data price) but less than at the indicator level (e.g. poor mobile coverage is unlikely to be compensated with high network quality). With this in mind, we have adopted the following aggregation rules:

— Dimension aggregation – arithmetic

— Enabler aggregation – arithmetic

— Index aggregation – geometric.

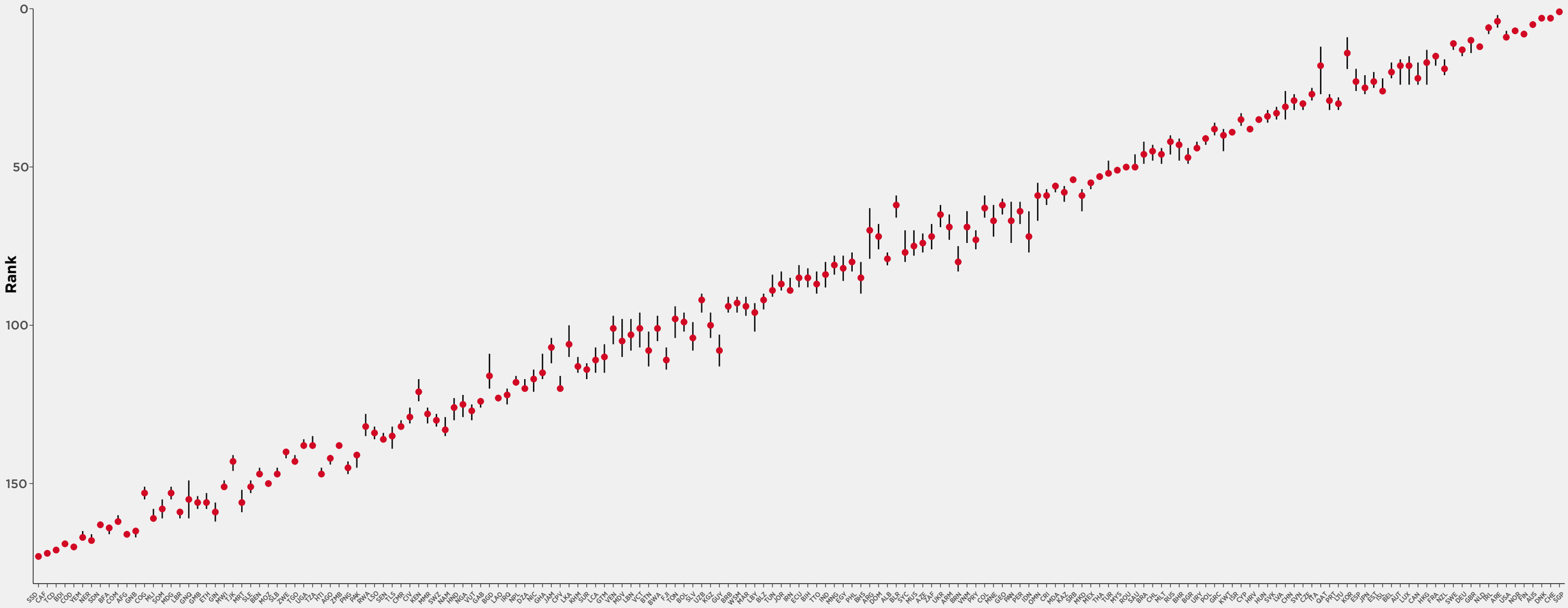
7. Uncertainty analysis

The JRC audit emphasised the importance of carrying out an uncertainty analysis of the MCI. We have therefore performed such an analysis with the updated MCI focusing on indicator, dimension and enabler weights.[24] The objective is to quantify the uncertainty in the ranks and scores of the MCI by randomly varying the weights assigned to each indicator, dimension and enabler. We rebuilt the 2023 index 500 times, each time randomly perturbing these weights by 25%. We then retrieved the median and 5–95% confidence intervals of the rankings and of the MCI scores. These are presented in Figures 1 and 2 below.

The conclusion of the uncertainty analysis is that uncertainty in the rankings is manageable, which is consistent with the JRC's findings. All countries fall within a 15 rankings precision interval and the median width of the 5–95% confidence interval is five rankings. This uncertainty analysis should be used to guide the conclusions drawn from the MCI. In particular, similar to the JRC, we recommend that a difference of two or three country rankings should not be taken as significant and a difference of 10 rankings or more should be taken as significant.

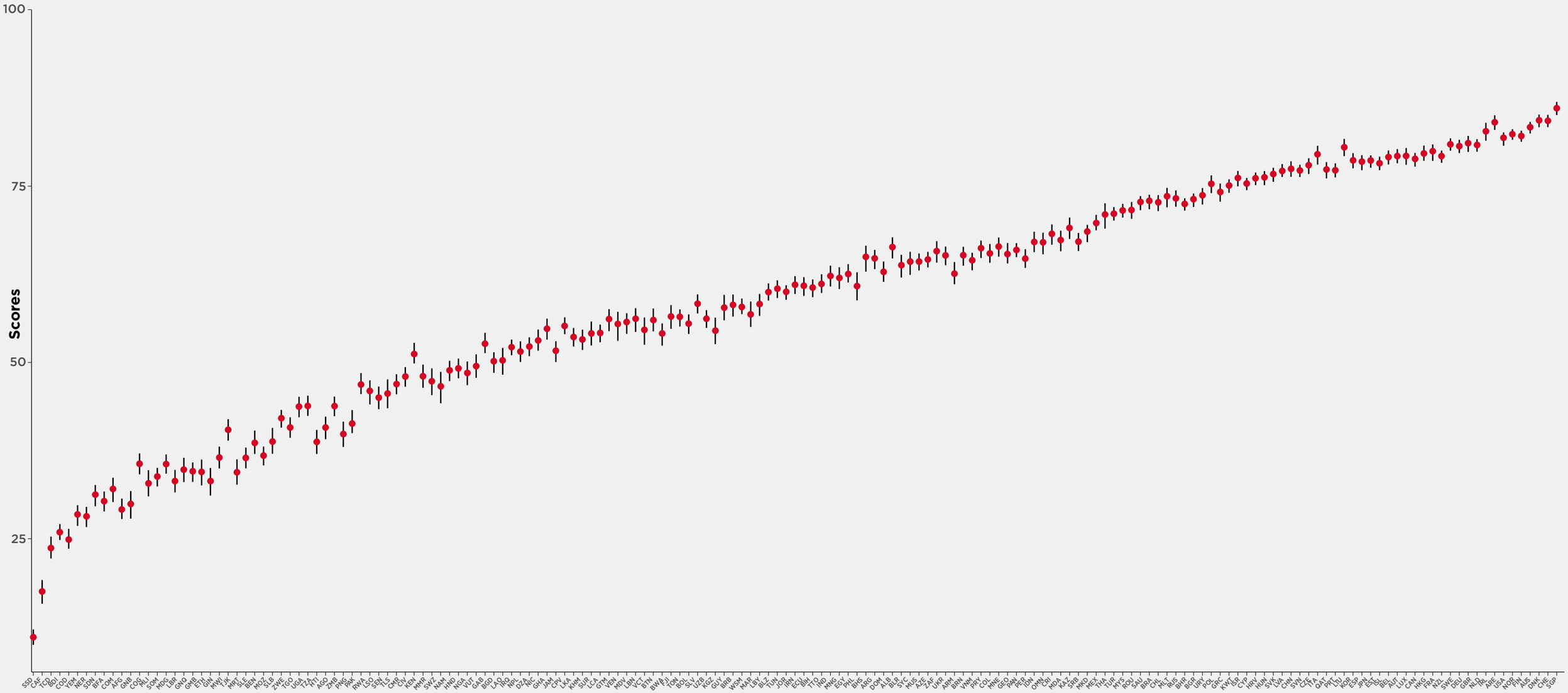
23. The analysis was carried out using the built-in function of the JRC "CoinR" R package.

Figure 1
Median ranking and 5–95% confidence intervals in the MCI rankings



Source: GSMA Intelligence

Figure 2
Median index score and 5–95% confidence intervals in the MCI index scores



Source: GSMA Intelligence

8. Feedback

Mobile internet connectivity is driving economic, societal and individual development around the world. In many locations, mobile technology provides the only opportunity for many to get online – and therefore the only way for millions of people to benefit from the transformational potential that the internet can deliver.

Through measuring and tracking the enablers of mobile internet connectivity, the MCI aims to support and accelerate the mobile industry's efforts to drive mobile internet connectivity and accelerate digital inclusion. With this in mind, the GSMA welcomes any feedback regarding the MCI. This includes hearing about how the MCI is being used by others in the mobile sector, and beyond. Please contact the Connected Society team (connectedsociety@gsma.com) with any questions, comments, suggestions or citations for the MCI.

GSMA Head Office
1 Angel Lane
London
EC4R 3AB
United Kingdom
Tel: +44 (0)20 7356 0600

