



20
50

connext
SOCIAL FUTURING CENTER

PÁL BÓDAY

METHODOLOGY USED TO COMPILE THE SOCIAL FUTURING INDEX

WORKING PAPER SERIES NO. 10/2020



Series Editors:

Petra Aczél
János Csák
Péter Szabadhegy
Zoltán Oszkár Szántó

Publisher:

Corvinus University of
Budapest,
Social Futuring Center,
Budapest 1093
Fővám tér 8.

ISBN 978-963-503-859-6
ISSN 2560-2357

Online version:

www.socialfuturing.com



1. METHODOLOGY USED TO COMPILE THE SFI¹

The SFI is a composite index of sub-indexes comprising a hierarchical indicator system based on the conceptual framework defined by the Social Futuring initiative. Simply put, the SFI is a weighted average of carefully selected indicators, which best capture the elements of Social Futuring.

The SFI comprises 28 indicators which were selected with the assistance of an expert panel. All indicators are normalized – after outliers were handled – on a scale of 0 to 100. The indicators are weighted and aggregated according to the structure of the SFI framework.

In order to best grasp and convey the concept of the indicator, a hierarchical structure was selected from a number of indicator system structures. The hierarchical structure makes it possible to create sub-indicators at different levels to examine the contexts of the conceptual framework, which makes the analysis even deeper. In general, such indicator systems are the most suitable for better presentation of complex, multi-dimensional phenomena.

In order to connect the normative standards with the pillars defined in the overall framework, definitions were prepared to describe the phenomena of nine essential paired intersections of the two aspects, based on which appropriate indicators could be selected.

1.1. SELECTING VARIABLES

An expert panel with specialists of different academic disciplines and statistics selected the indicators and compiled the first set of indicators that best suited the written definitions. The selection process of the indicators followed the basic principle that indicators had to:

- be measurable/available,

¹ All data and methods used during the compilation of the SFI 2020 are available in detail on the SFI website in order to make the methodological procedures replicable.

- be accessible from official, publicly available sources
- have at least OECD-country coverage,
be without or have limited overlapping with other indicators,
and
- have a measurable range of the indicator.

Several workshops served to finalize and fine tune the indicator set to avoid overlaps as well as to maintain a balance between the different elements of the framework. The first set covered around 120 indicators, which was reduced to the final 28 essential indicators, which are deemed relevant and meet the above-mentioned basic principles.

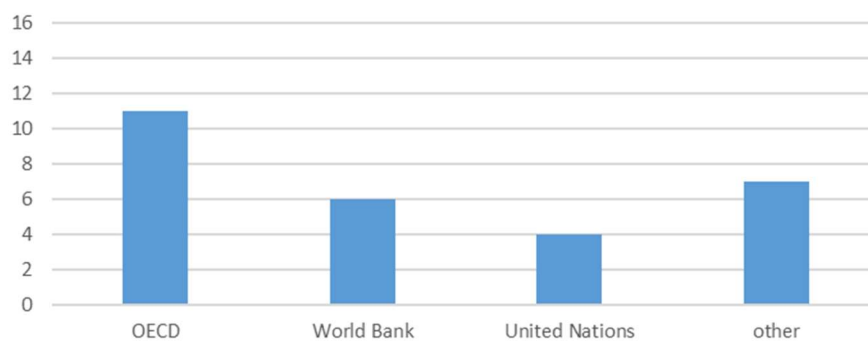


Figure 1: Sources of indicators

There are three types of indicators that have been chosen:

1. Relative indicators: relative indicators are obtained by dividing an indicator by another indicator- in order to maintain comparability between countries. The basic indicators used in the denominator are GDP, population or others such as the number of households or area.
2. Scales: some indicators are defined to be measured along a predefined range.
3. Product (or mix) indicators: to measure both temporal change and the current level of a given phenomena. The indicator is a product of two basic indicators: the percentage change in the phenomenon over time and the percentage deviation of the current value from the average.

For each indicator, the most recent data available was used. (Available until 1st May 2020). In most cases, 2017-2018 data were available. In some cases, the model relies on earlier data.

For each indicator, the direction (positive or negative) was determined to the concept of social futuring, based on its relevance.

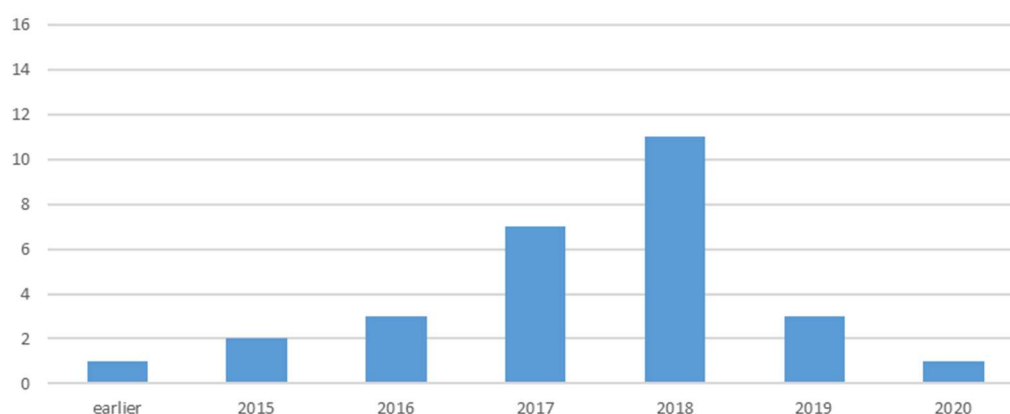


Figure 2: Reference years of indicators

1.2. IMPUTATION

Although the selection of indicators was based on maximum country coverage, in the case of some indicators, data from a few countries were either missing or significantly different (4-5 years) in time from other countries. In these cases, the data were usually imputed using other reliable sources or in some cases supplemented with the value of a similar country. The replaced data represent only 2.5 percent of the total data used.

1.3. HANDLING OUTLIERS

Outliers are individual values that fall outside of the overall pattern of a data set. Outliers were filtered out before data were normalized, since outliers could significantly affect normalized values, especially when applying the min-max approach. The interquartile range rule was used for detecting the presence of outliers. The interquartile

range (IQR) is calculated by subtracting the first quartile (Q1) from the third quartile (Q3). According to the common rule if an individual value is higher than $Q3 + 1.5 \cdot IQR$ or smaller than $Q1 - 1.5 \cdot IQR$, the data is considered as an outlier. Outliers are replaced with $Q3 + 1.5 \cdot IQR$ or $Q1 - 1.5 \cdot IQR$.

1.4. NORMALIZATION

Normalization is required prior to any data aggregation, as the indicators in a data set often have different measurement units or order of magnitude. Different normalization and aggregation techniques were tested (min-max, standardization, ranking, above-below mean, categories). The min-max method was chosen because it best met the needs of the model, the compilation of the hierarchical composite indicator. There are no negative numbers, or there is no problem with handling 0, additivity is retained.

1.5. WEIGHTING AND AGGREGATION

Weights were determined by expert consensus. They were defined on the basis of the conceptual framework, taking into account the importance of normative standards. Within the normative standards, two dimensions (Assets and Family) were given higher weights within its normative standard. All indicators within each dimension were given equal weights.

Normative standards	Weights by normative standard	Dimensions	Weights by dimension	Number of indicators within dimension	Weights by indicator
Peace & Security	40	Defense & Safety	10	3 indicators	3,33
		Assets	20	4 indicators	5,00
		Functionality	10	3 indicators	3,33
Attachment	30	Patriotism	7,5	2 indicators	3,75
		Family	15	3 indicators	5,00
		Spirituality	7,5	2 indicators	3,75
Care	20	Self-reliance	10	3 indicators	3,33
		Material Advancement	10	3 indicators	3,33
Balance	10	Wellbeing & Generativity	10	5 indicators	2,00

*Table 1:
Weighting of the components of the SFI*

Aggregation was based on weights and normalized indicator values. The final SFI and/or any sub-indicator is the weighted sum of the normalized indicator values. Also, the composite indicator at any given level (dimension or normative standard) can be built from the sub-indicators that make it up. This greatly facilitates the analysis of the effect of the indicator composition.

Composite indicators can be interpreted as the weighted sum of the normalized indicator values (this makes it possible to examine the weight of sub-indicators within higher-level indicators), or on a scale from 0 to 100.

1.6. CLUSTERING

The data were analyzed and compared using several methods for the cluster analysis.

For clustering we used the k-means algorithm, which is one of the most popular clustering algorithms. In the k-means algorithm, a set of data is classified using a certain number of clusters (k clusters) which are initialized apriori. It defines k centroids, one for each cluster and then considers data objects belonging to the given data set and associates these data objects to the closest centroid. Euclidean distance is considered to determine the distance between data objects and the centroids.

To examine the relationship and similarity of the countries, we calculated the clusters for clusters between 2 and 10 at each indicator level (indicator, dimension, normative standard).

2. INDICATORS

2.1. INDICATORS USED FOR PEACE & SECURITY NORMATIVE STANDARD – DEFENSE & SAFETY DIMENSION

(reactive, policy sensitivity: yes)

1. Political stability and absence of violence or terrorism (direction: positive, weight: 3,33%)

Definition: Political stability and absence of violence or terrorism measures perceptions of the likelihood of political instability and/or politically-motivated violence, including terrorism. Estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, i.e. ranging from approximately -2.5 to 2.5.

Unit of measure: index (-2,5 to 2,5)

Source of data: WB, <http://info.worldbank.org/governance/wgi/Home/Reports>

2. Robbery (direction: negative, weight: 3,33%)

Definition: Robbery is a property crime that involves the use of violence or threat of violence. Theft of property from a person, overcoming resistance by force or threat of force. Robbery included muggings, bagsnatching and theft with violence.

Unit of measure: per 100,000 population

Source of data: United Nations Office on Drugs and Crime (UNODC) <https://dataunodc.un.org/crime/robbery>

3. Military expenditure (direction: positive, weight: 3,33%)

Definition: Military expenditures data from SIPRI are derived from the NATO definition, which includes all current and capital expenditures on the armed forces, including peacekeeping forces; defense ministries and other government agencies engaged in defense projects; paramilitary forces, if these are

judged to be trained and equipped for military operations; and military space activities.

Unit of measure: percent of GDP

Source of data: WB, <https://data.worldbank.org/indicator/MS.MIL.XPND.GD.ZS>

2.2. INDICATORS USED FOR PEACE & SECURITY NORMATIVE STANDARD– ASSETS DIMENSION

(proactive, policy sensitivity: no)

4. Ecological balance (direction: positive, weight: 5%)

Definition: The difference between a population's Ecological Footprint and a country's biocapacity. If a country's demand exceeds its biocapacity, it has an ecological deficit. If a country's biocapacity exceeds its Ecological Footprint, it has an ecological reserve.

Unit of measure: global hectare

Source of data: Global Footprint Network, <http://data.footprintnetwork.org/#/exploreData>

5. Arable land (direction: positive, weight: 5%)

Definition: Arable land (hectares per person) includes land defined by the FAO as land under temporary crops (double-cropped areas are counted once), temporary meadows for mowing or for pasture, land under market or kitchen gardens, and land temporarily fallow. Land abandoned as a result of shifting cultivation is excluded.

Unit of measure: hectares per person

Source of data: WB, <http://wdi.worldbank.org/table/3.1#>

6. Net energy imports (direction: negative, weight: 5%)

Definition: Net energy imports are estimated as energy use less production, both measured in oil equivalents.

Unit of measure: percent of energy use

Source of data: WB, <http://wdi.worldbank.org/table/3.8>

7. Renewable water resources (direction: positive, weight: 5%)

Definition: Total annual actual renewable water resources per inhabitant [Total renewable water resources per capita] = [Total renewable water resources]*1000000/[Total population].

Unit of measure: cubic meter/inhabitant/year

Source of data: FAO, <http://www.fao.org/nr/water/aquastat/data/query/index.html>

2.3. INDICATORS USED FOR PEACE & SECURITY NORMATIVE STANDARD – FUNCTIONALITY DIMENSION

(active, policy sensitivity: yes)

8. High-technology exports (direction: positive, weight: 3,33%)

Definition: High-technology exports are products with high R&D intensity, such as in aerospace, computers, pharmaceuticals, scientific instruments, and electrical machinery. (Data are in percentage of manufactured exports). Because industrial sectors specializing in a few high-technology products may also produce low-technology products, the product approach is more appropriate for international trade.

Unit of measure: percent of manufactured exports

Source of data: WB, <https://data.worldbank.org/indicator/TX.VAL.TECH.MF.ZS>

9. Road density (per capita) (direction: positive, weight: 3,33%)

Definition: Road density is the ratio of the length of the country's total road network to the country's population. The road network includes all roads in the country: motorways, highways, main or national roads, secondary or regional roads, and other urban and rural roads. The Global Roads Inventory Project is a harmonized global dataset of approximately 60 geospatial datasets on road infrastructure. The resulting dataset covers 222 countries and includes over 21 million km of roads, which is two to three times the total length in the currently best available country-based global roads datasets.

Unit of measure: km per capita

Source of data: Global Roads Inventory Project + own calculation, https://stats.oecd.org/Index.aspx?DataSetCode=ITF_INDICATORS

10. Households broadband internet connection (direction: positive, weight: 3,33%)

Definition: Household broadband access provides a measure of the uptake of broadband technology by households. It refers to the share of households that have purchased subscriptions to fixed-line or mobile broadband services.

Unit of measure: percent of households

Source of data: OECD, <https://goingdigital.oecd.org/en/indicator/13/>

2.4. INDICATORS USED FOR ATTACHMENT NORMATIVE STANDARD – PATRIOTISM DIMENSION

(active, policy sensitivity: yes)

11. Persons living abroad (direction: negative, weight: 3,75%)

Definition: Ratio of the estimates of the international migrant (mid-year) stock, by origin and the total mid-year population (obtained from World Population Prospects: The 2017 Revision).

Unit of measure: percent of population of origin country

Source of data: UN, <https://www.un.org/en/development/desa/population/migration/data/estimates2/estimates19.asp>

12. Registered voters who actually voted (direction: positive, weight: 3,75%)

Definition: The total number of votes cast (valid or invalid) divided by the number of names on the voters' register, expressed as a percentage. Parliamentary Elections: The parliamentary elections displayed in the Voter Turnout database are elections to the national legislative body of a country or territory. In case the legislative body has two chambers, only the second (lower) chamber is included. If elections are carried out in two rounds (using the Two-Round System TRS), only the second election round is included.

Unit of measure: percent

Source of data: IDEA, <https://www.idea.int/data-tools/question-view/521>

2.5. INDICATORS USED FOR ATTACHMENT NORMATIVE STANDARD – FAMILY DIMENSION

(active, policy sensitivity: yes)

13. Employees working very long hours – work-life balance (direction: negative, weight: 5%)

Definition: Percentage of all employees usually working 50 hours or more per week.

Unit of measure: percent

Source of data: OECD, <https://stats.oecd.org/Index.aspx?DataSetCode=BLI#>

14. Value of family benefits (direction: positive, weight: 5%)

Definition: Total family benefits for a two-parent, two-earner family for 2 children with a youngest child aged 6, as a % of average full-time earnings.

Unit of measure: percent of average full-time earnings

Source of data: OECD, <https://stats.oecd.org/Index.aspx?QueryId=79865#>

15. Single person households (direction: negative, weight: 5%)

Definition: Share of single person households among all households.

Unit of measure: percent

Source of data: Eurostat, http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=ilc_lvph02&lang=en

2.6. INDICATORS USED FOR ATTACHMENT NORMATIVE STANDARD – SPIRITUALITY DIMENSION

(proactive, policy sensitivity: no)

- 16.** Important to follow traditions and customs (direction: negative, weight: 3,75%)

Definition: On a scale from 1 to 6, where 1 means very much like me and 6 means not at all like me.

Unit of measure: scale 1 to 6

Source of data: World Values Survey, <http://www.worldvaluessurvey.org/WVSDocumentationWV6.jsp>

- 17.** Self-reported religiousness (direction: positive, weight: 3,75%)

Definition: The share of those, who replied as religious person for the question. You are: (1) A religious person, (2) Not a religious person, (3) A convinced atheist.

Unit of measure: percent

Source of data: World Values Survey, <http://www.worldvaluessurvey.org/WVSONline.jsp> – <http://www.europeansocialsurvey.org/>

2.7. INDICATORS USED FOR CARE NORMATIVE STANDARD – SELF-RELIANCE DIMENSION

(proactive, policy sensitivity: yes)

- 18.** Mean years of schooling (direction: positive, weight: 3,33%)

Definition: Average number of years of education received by people ages 25 and older, converted from education attainment levels using official durations of each level.

Unit of measure: years

Source of data: UNDP, <http://hdr.undp.org/en/indicators/103006>

19. Unemployment rate (direction: negative, weight: 3,33%)

Definition: Unemployment rate is the number of unemployed people as a percentage of the labor force, where the latter consists of the unemployed plus those in paid or self-employment. Unemployed people are those who report that they are without work, that they are available for work and that they have taken active steps to find work in the last four weeks.

Unit of measure: percent

Source of data: OECD, https://stats.oecd.org/Index.aspx?DataSetCode=LFS_SEXAGE_I_R

20. Life expectancy (mix) (direction: positive, weight: 3,33%)

Definition: Life expectancy at birth is defined as how long, on average, a newborn can expect to live, if current death rates do not change. The indicator is calculated as the product of the long term change (2010 to 2017) and the distance to maximum of the current value.

Unit of measure: percent

Source of data: OECD, https://stats.oecd.org/sdmx-json/data/DP_LIVE/.LIFEEXP.../OECD?contentType=csv&detail=code&separator=comma&csv-lang=en

2.8. INDICATORS USED FOR CARE NORMATIVE STANDARD – MATERIAL ADVANCEMENT DIMENSION

(active, policy sensitivity: no)

21. Household expenditure (direction: positive, weight: 3,33%)

Definition: Household spending is the amount of final consumption expenditure made by resident households to meet their everyday needs, such as food, clothing, housing (rent), energy, transport, durable goods (notably cars), health costs, leisure, and miscellaneous services. The indicator shows their expenditure relative to the GDP.

Unit of measure: percent of GDP

Source of data: OECD, https://stats.oecd.org/Index.aspx?DataSetCode=SNA_TABLE5

22. Child relative income poverty rate (direction: negative, weight: 3,33%)

Definition: The percentage of children (0-17 year-olds) with an equivalized household disposable income (i.e. an income after taxes and transfers adjusted for household size) below the poverty threshold. The poverty threshold is set here at 50% of the median disposable income in each country.

Unit of measure: percent of population 0-17 years

Source of data: OECD, http://www.oecd.org/els/soc/CO_2_2_Child_Poverty.xlsx

23. GDP/capita (mix) (direction: positive, weight: 3,33%)

Definition: Gross domestic product (GDP) is the standard measure of the value added created through the production of goods and services in a country during a certain period. The indicator is calculated as the product of the long term change (2010 to 2017) and the distance to the OECD average of the current value in USD.

Unit of measure: percent

Source of data: OECD, https://stats.oecd.org/Index.aspx?DataSetCode=SNA_TABLE1

2.9. INDICATORS USED FOR BALANCE NORMATIVE STANDARD – WELLBEING & GENERATIVITY DIMENSION

(proactive, policy sensitivity: no)

24. Transition of educational attainment level from parents to current adults (direction: positive, weight: 2%)

Definition: Transitions from previous generation – from pre-primary, primary and lower secondary education of parents to tertiary education.

Unit of measure: percent

Source of data: Eurostat, http://appsso.eurostat.ec.europa.eu/nui/show.do?lang=en&dataset=ilc_igtp01

25. Fertility (mix) (direction: positive, weight: 2%)

Definition: The total fertility rate in a specific year is defined as the total number of children that would be born to each woman if she were to live to the end of her child-bearing years and give birth to children in alignment with the prevailing age-specific fertility rates. The indicator is calculated as the product of the long term change (2010 to 2017) and the distance to the OECD average of the current value.

Unit of measure: percent

Source of data: OECD, https://stats.oecd.org/viewhtml.aspx?datasetcode=HEALTH_DEMR&lang=en#

26. Age dependency (direction: negative, weight: 2%)

Definition: The ratio of dependents (people younger than 15 or older than 64) to the working-age population (15-64).

Unit of measure: percent of working-age population

Source of data: WB, <https://data.worldbank.org/indicator/SP.POP.DPND>

27. Anti-depressant usage (direction: negative, weight: 2%)

Definition: Antidepressant drugs consumption in DDD. Defined daily dose (DDD) is the assumed average maintenance dose per day for a drug used for its main indication in adults.

Unit of measure: Defined daily dosage per 1 000 people per day

Source of data: OECD, Health statistics, <http://dx.doi.org/10.1787/888933605540>

28. Gini-coefficient (income) (direction: negative, weight: 2%)

Definition: Gini index measures the extent to which the distribution of income (or, in some cases, consumption expenditure) among individuals or households within an economy deviates from a perfectly equal distribution. A Lorenz curve plots the cumulative percentages of total income received against the cumulative number of recipients, starting with the poorest individual or household. The Gini index measures the area between the Lorenz curve and a hypothetical line of absolute equality, expressed as a percentage of the maximum area under the line. Thus, a Gini index of 0 represents perfect equality, while an index of 100 implies perfect inequality.

Unit of measure: 0-100

Source of data: OECD, <https://data.oecd.org/inequality/income-inequality.htm>

NOTES

NOTES

NOTES

20
50



Social Futuring Center (SFC) is an independent and multidisciplinary research unit of the Corvinus University of Budapest (CUB). Our aims are to develop the conceptual and normative framework of social futuring, to construct the Social Futuring Index (SFI) and to manage the **ConNext2050** research project. The main scope of its research is the analysis and interpretation of social futuring of different social entities, focusing on short and long-term future changes (2017-2050).

The SFC periodically publishes **working papers** that highlight the findings of its research. They are published to stimulate discussion and contribute to the advancement of our knowledge of multidisciplinary matters related to philosophy, sociology, psychology, bionics, informatics, economics, political science, environmental studies, futures studies, network science. SFC working papers are available online on the **www.socialfuturing.com** website.

