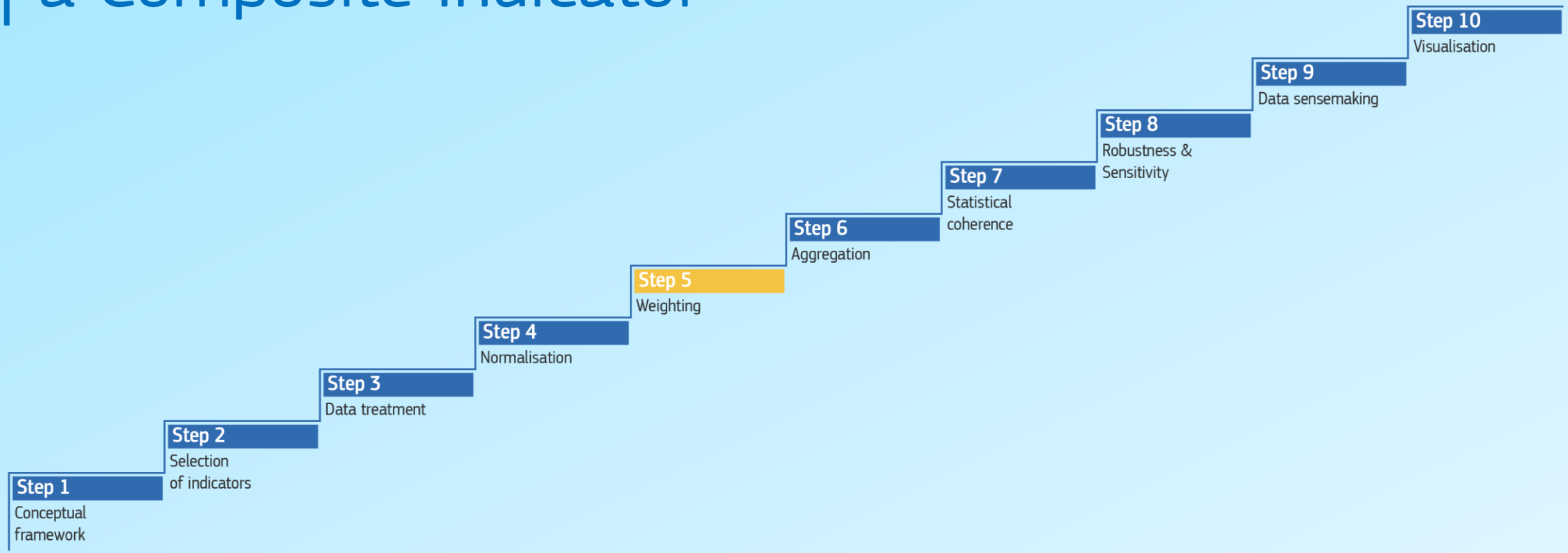
Step 5
Weighting methods

18th JRC Annual training on Composite Indicators and Scoreboards

Matija Kovacic

10 STEPS to build a Composite Indicator



European Commission

Weighting in composite indicators

- **Meaning of weighting**:
  - Relative importance of indicators/pillars/sub-pillars ("*explicit importance*")
  - Trade-off between indicators/pillars/sub-pillars ("*implicit importance*")*****
- **Selecting a weighting scheme** is not a simple task:
  - There is no "*one-size-fits-all*" solution
  - Weight may have a significant impact on final scores and ranks (the so-called "*index problem*")
  - Stakeholders may have different opinions on choosing weighting scheme
- **A composite is your own product**: theoretical framework + transparency

Approaches to setting weights

- **Equal weights**
- **Weighting based on statistical methods**
  - Principal component analysis/Factor analysis
  - Data envelopment analysis
  - Regression approach
- **Weights based on public/expert opinion**
  - Budget allocation and Analytic hierarchy process
  - Conjoint analysis

Approaches to setting weights: examples

Composite Indicator	Weighting scheme
Human Development Index	Equal weights
Multidimensional Poverty Index	Equal weights Expert opinion
Quality of Life Index	Equal weights
Better Life Index	Equal weights
Social Progress Index	Principal component analysis
Corruption Perception Index	Equal weights
Technological Achievement Index	Expert opinion
The Cultural and Creative Cities Monitor	Expert opinion
Rule of Law Index	Equal weights Expert opinion
Environmental Performance Index	Expert opinion
Gender Equality Index	Expert opinion
ICT 2015 Development Index	Principal component analysis

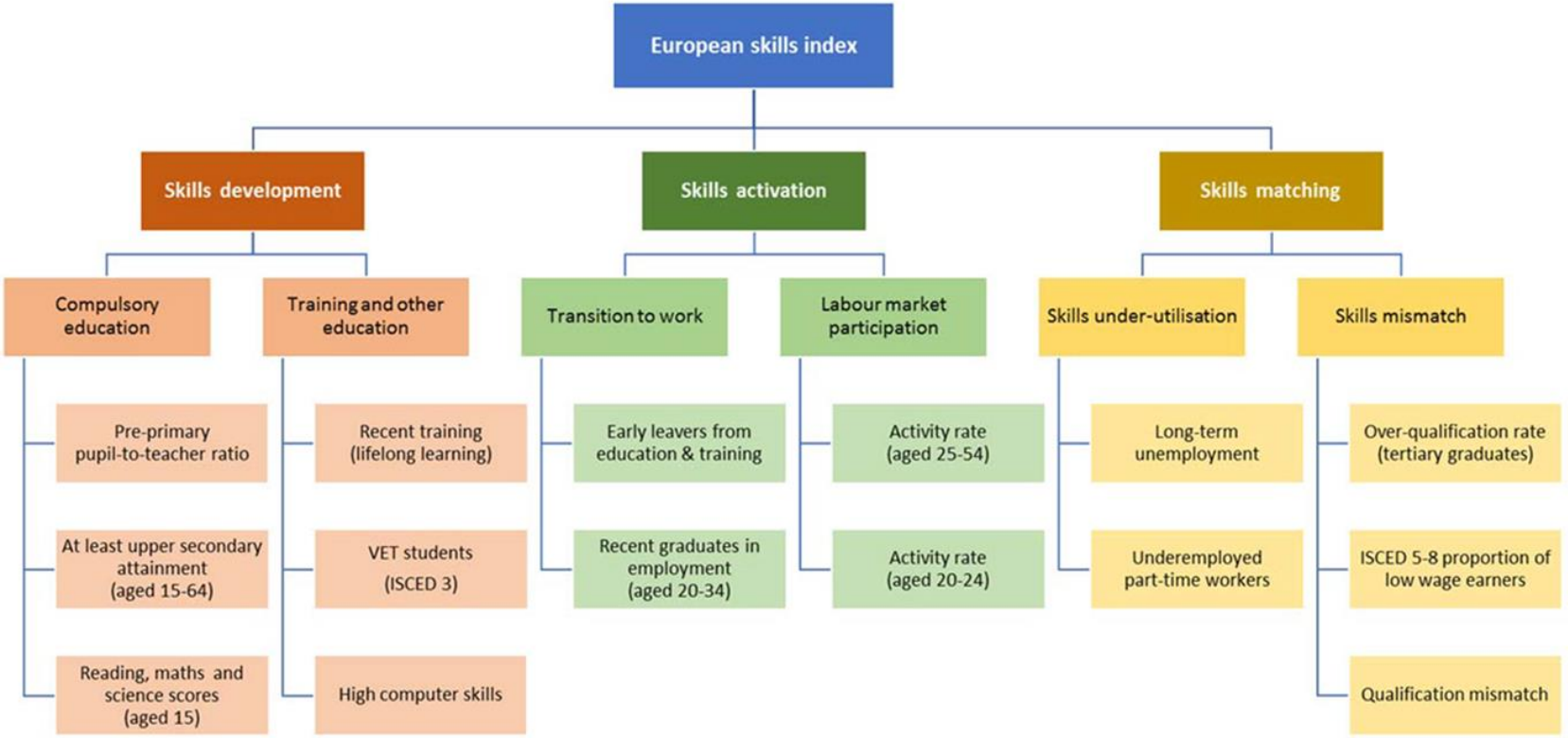
Equal weighting

- Equal weighting is *the most common* scheme appearing in composite indicators
- **Justifications** of choosing equal weights:
  - *Simple and easy to communicate*
  - *No à priori knowledge and no clear reference about the importance of the elements*
  - *No agreement between stakeholders*
  - *"Problems" with statistical approaches*

However, equal weighting …

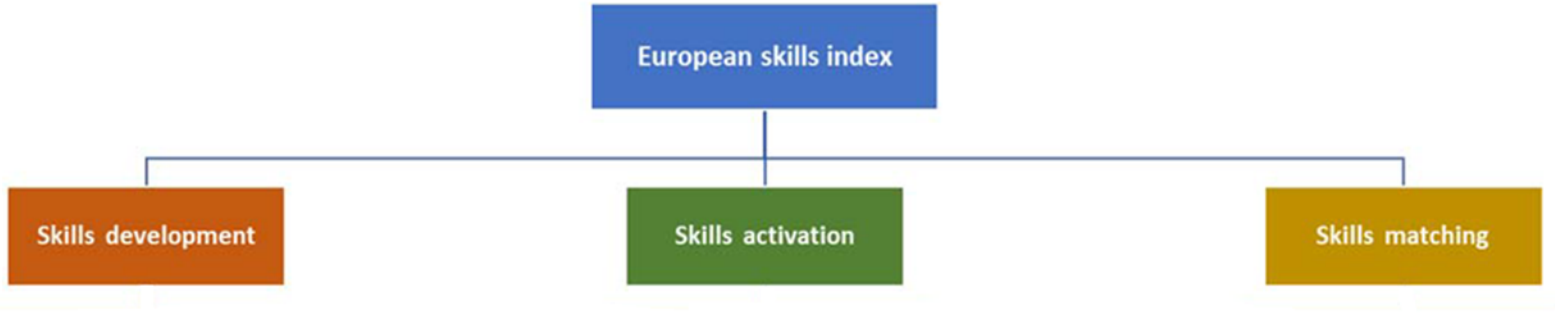
- **<u>does not mean</u> not distributing weights at all**
- **<u>does not mean</u> equal "contribution" of the indicators to the composite**

Example: European Skills Index



Source: *European Skills Index (2020), Cedefop.*

Example: European Skills Index



Equal Weights

Pillars	Pearson Correlation Coefficient	R^2
Skills Development	0.80	0.64
Skills Activation	0.81	0.66
Skills Matching	0.64	0.41

Adjusted Weights

Pillars	Pearson Correlation Coefficient	R^2
Skills Development	0.77	0.59
Skills Activation	0.76	0.58
Skills Matching	0.71	0.50

Principal component analysis (PCA)

- **Objective**: find a set of weights which, combined with the corresponding indicators, explains **most of the variance**

- These combinations (averages) are called **principal components (PC)**:

$$PC_1 = c_{11}X_1 + c_{21}X_2 + \ldots + c_{n1}X_n$$
$$PC_2 = c_{12}X_1 + c_{22}X_2 + \ldots + c_{n2}X_n$$
$$\ldots$$

  - Each principal component is a new variable computed as a linear combination of the original (standardized) variables

- Choose the component (*i.e.*, combination of coefficients) that explains **most of the observed variance**

How should we choose?

Several methods exist. The **3** most common are:

Kaiser–Guttman "Eigenvalues greater than one" criterion (Guttman,1954; Kaiser, 1960): select all components with eigenvalues greater than 1 (or 0.9)

Certain percentage of explained variance**:** for example **>60**, **70**%, 80%,…

Cattell's scree test (Cattell, 1966) "*Above the elbow*" approach

- **Objective:** one Principal Component (one-dim solution) – single latent phenomenon

Weights based on PCA

- The coefficients (standardised) of the first principal component are used as weights:

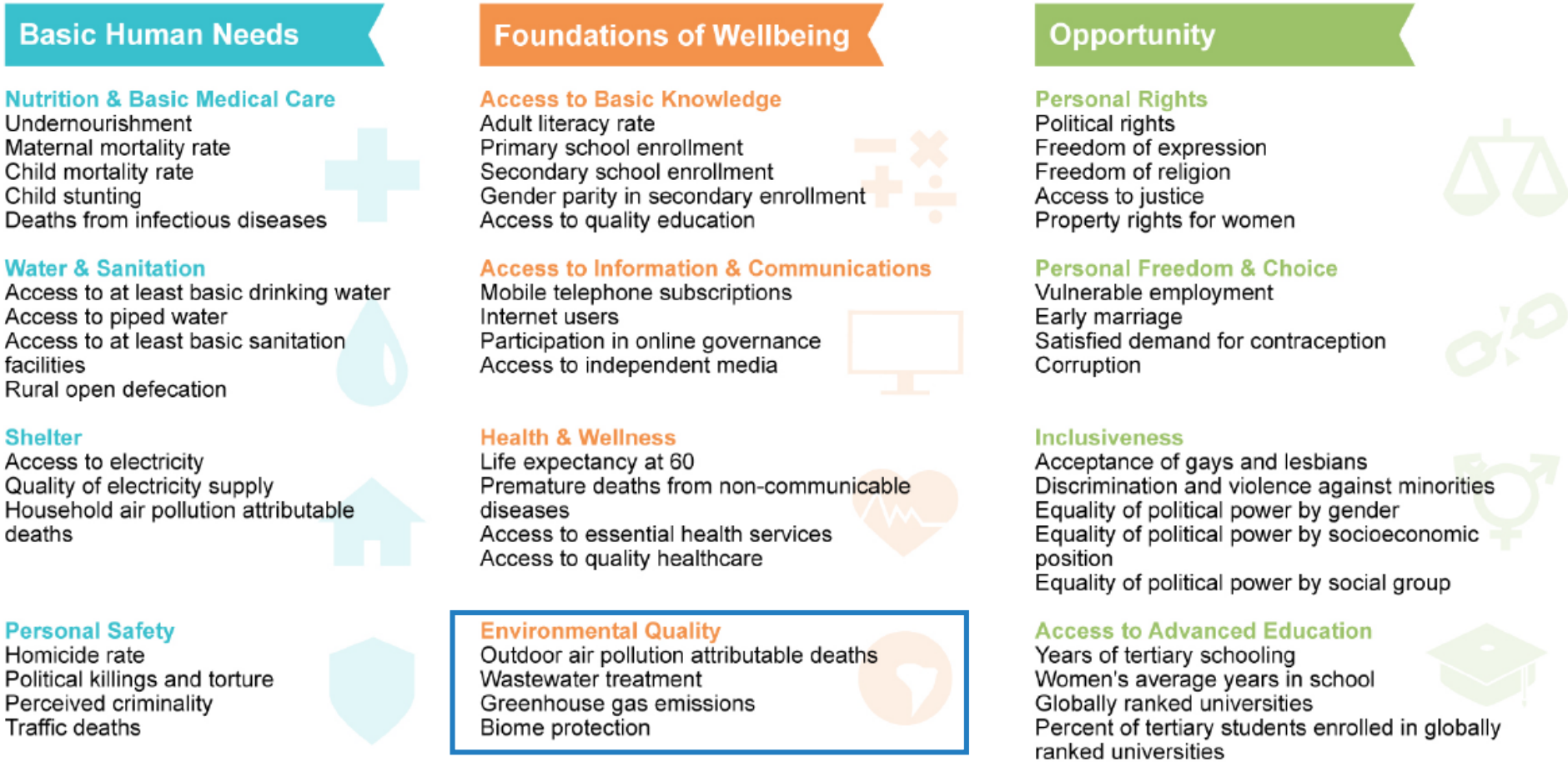
$$PC_{1*} = c_1X_1 + c_2X_2 + \ldots + c_nX_n$$

$$\downarrow \qquad \downarrow \qquad \qquad \downarrow$$

$$w_1 \qquad w_2 \qquad \ldots \qquad w_n$$

- **Positive aspects:**
  - Empirical and objective option for weight selection,
  - Maximizes the variance of the units.
- **Potential disadvantages**:
  - Assigns lower weights to dimensions that are poorly correlated,
  - Less straightforward and transparent.

Example 1: Social Progress Index (SPI)



- **Illustrative example**: PCA analysis for the indicators within the *Environmental Quality* component

Example 1: Social Progress Index (SPI)

Total variance explained

Component	Eigenvalue	Variance (%)	Cumulative variance (%)
PC1	2.28	57.11	57.11
PC2	0.87	21.81	78.92
PC3	0.51	12.80	91.71
PC4	0.33	8.29	100

Pearson correlation coefficient between indicators and principal components

Indicators	PC1	PC2	PC3	PC4
Outdoor air pollution attributable deaths	0.88	-0.12	-0.11	-0.45
Wastewater treatment	0.83	-0.13	-0.44	0.32
Greenhouse gas emissions	0.78	-0.26	0.55	0.16
Biome protection	0.47	0.88	0.08	0.03

Normalize to unity

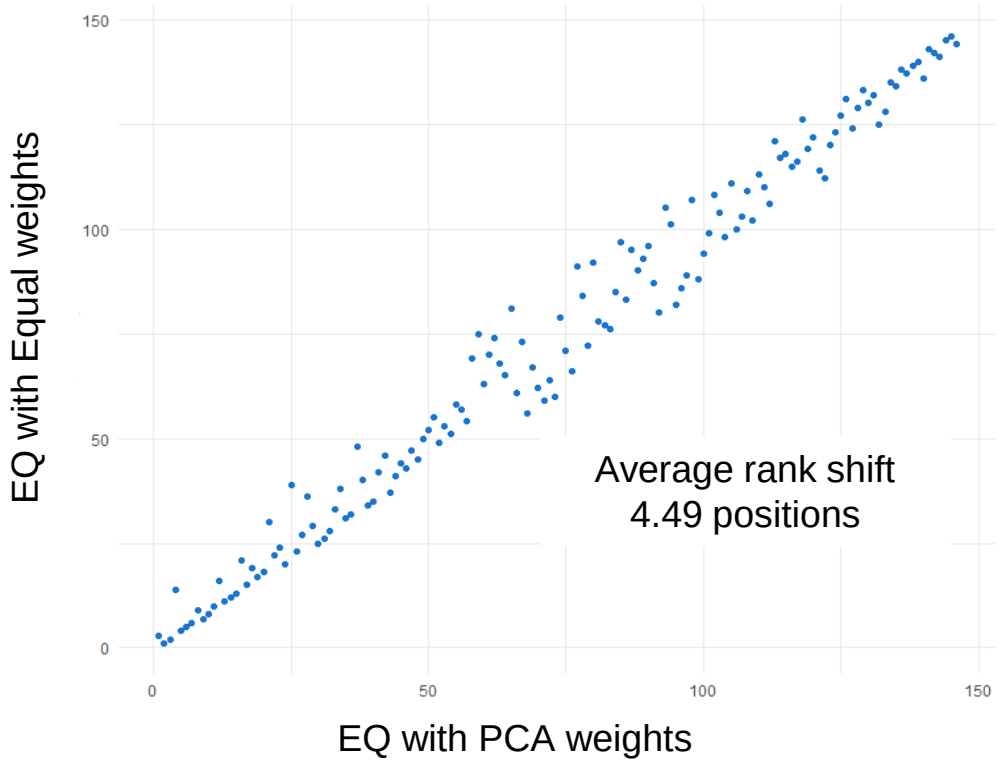
PC1*norm*
0.30
0.28
0.26
0.16

$$PC1 = 0.88\ X1 + 0.83\ X2 + 0.78\ X3 + 0.47\ X4$$

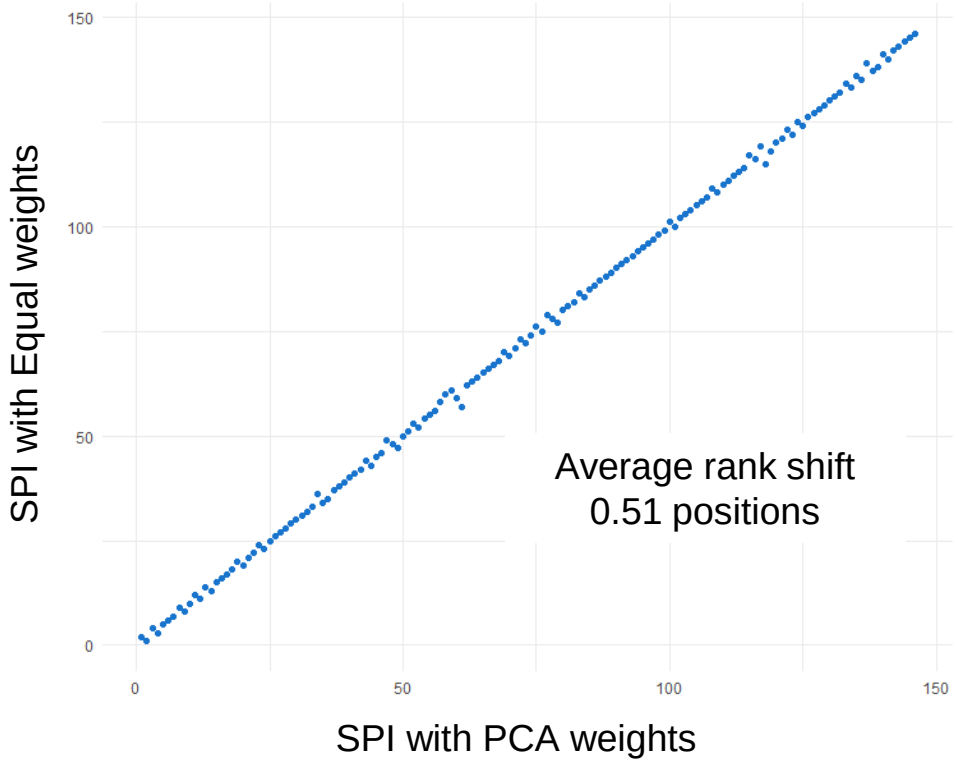
$$PC1_{norm} = 0.30\ X1 + 0.28\ X2 + 0.26\ X3 + 0.16\ X4$$

Example 1: Social Progress Index (SPI)

Comparing weighting schemes
Environmental Quality (EQ) Component



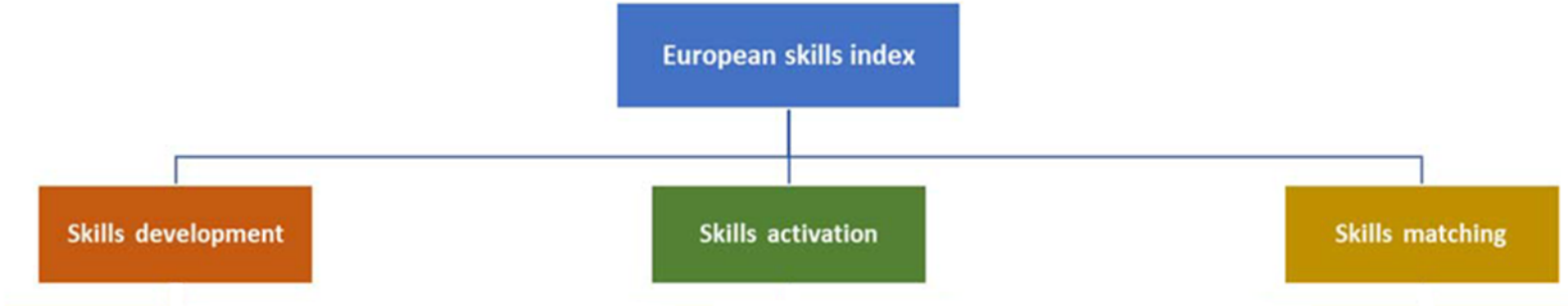
Comparing weighting schemes
SPI Index score



- **Conclusion**: very similar ranking (in this case both may work well) – does not mean they are equivalent

European Commission

Example 2: European Skills Index (SPI)



Total variance explained

Component	Eigenvalue	Variance	Cumulative variance
PC1	1.75	58	58
PC2	0.88	29	88
PC3	0.37	12	100

Pearson correlation coefficient

Pillars	PC1	PC2	PC3
Skills Development	0.88	-0.16	-0.44
Skills Activation	0.84	-0.36	0.41
Skills Matching	0.51	0.85	0.09

PC1*norm*
0.40
0.38
0.23

Example 2: European Skills Index (SPI)

Equal Weights

Pillars	Pearson Correlation Coefficient	R^2
Skills Development	0.80	0.64
Skills Activation	0.81	0.66
Skills Matching	0.64	0.41

PCA Weights

Pillars	Pearson Correlation Coefficient	R^2
Skills Development	0.84	0.70
Skills Activation	0.87	0.75
Skills Matching	0.52	0.27

- With PCA weighting the index becomes even **more unbalanced** with respect to equal weighting scheme
- PCA "*punishes*" the indicators which are poorly correlated with the others
- However, these indicators, from a conceptual point of view, may be very important and should have a higher weight => **conceptual framework**

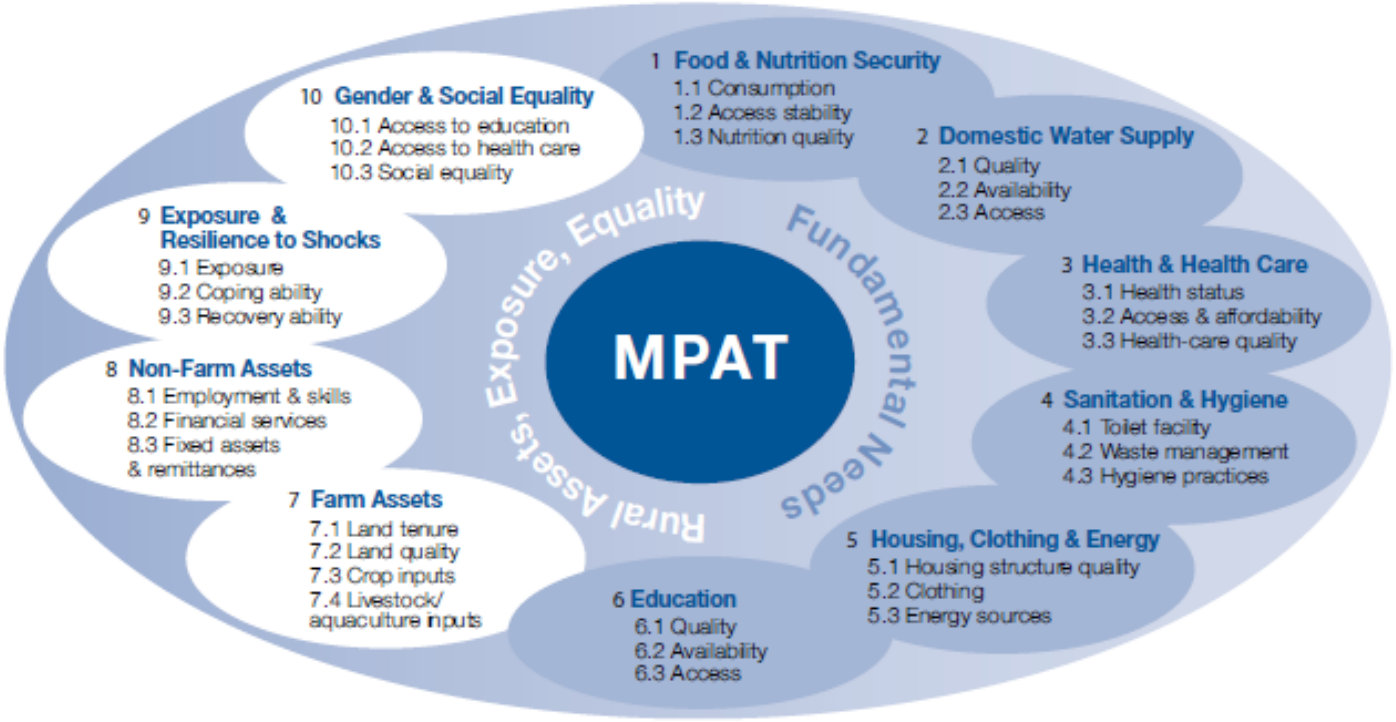
Expert opinion

- There are two typical methods to elicit views from experts: **budget allocation (BAP)** and **analytic hierarchy process (AHP)**

  - In **BAP** a panel of experts is given *N* points to distribute to the indicators, or groups of indicators (sub-pillars/pillars) - an **average** of the experts' choices is used as a weight
  - **Critical issues**: selection of experts (number and expertise) and the total number of indicators to be evaluated:

- Experts should **not** be specialists for individual indicators, rather for the given sub-index (larger "field" expertise)
- not more than **10** indicators – risk of "*circular thinking*"

BAP - Example 1: Multidimensional poverty assessment tool (MPAT)



Selection of Experts:

- **42** experts from 10 countries and 28 organizations (mainly UN agencies and universities)
- selection based on expertise on poverty assessment tools in developing countries

Purpose: Eliciting weights to be assigned to the subcomponents of each of the dimensions

BAP - Example 1: MPAT

Sub1.1	43%
Sub1.2	32%
Sub1.3	25%
Sub2.1	29%
Sub2.2	38%
Sub2.3	33%
Sub3.1	38%
Sub3.2	34%
Sub3.3	28%
Sub4.1	38%
Sub4.2	26%
Sub4.3	36%
Sub5.1	38%
Sub5.2	33%
Sub5.3	29%
Sub6.1	31%
Sub6.2	33%
Sub6.3	36%
Sub7.1	36%
Sub7.2	24%
Sub7.3	20%
Sub7.4	20%

Food and nutrition

Sub8.1	39%
Sub8.2	33%
Sub8.3	28%
Sub9.1	33%
Sub9.2	34%
Sub9.3	33%
Sub10.1	31%
Sub10.2	36%
Sub10.3	33%

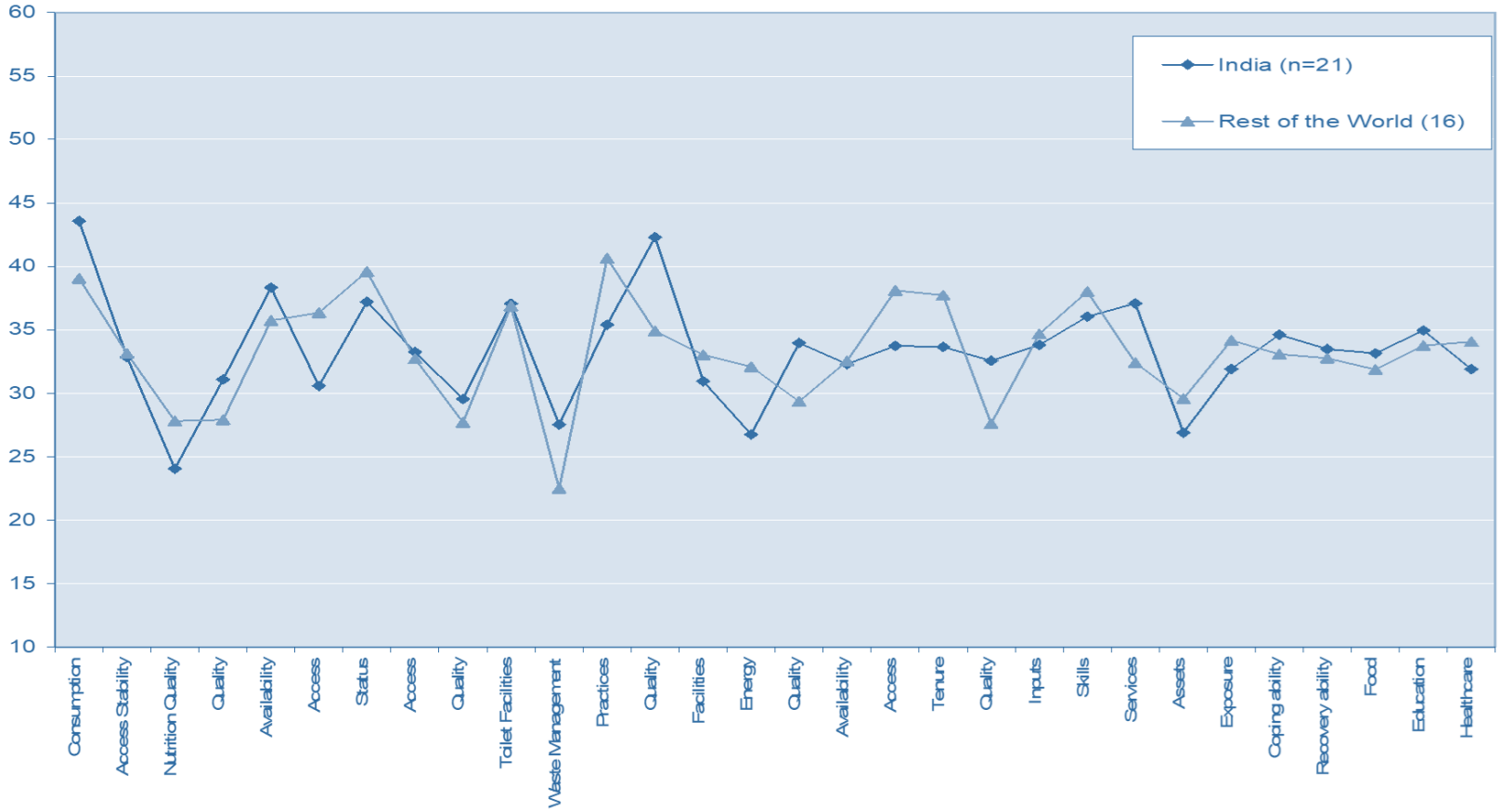
Exposure to shocks

Gender and social equality

Education

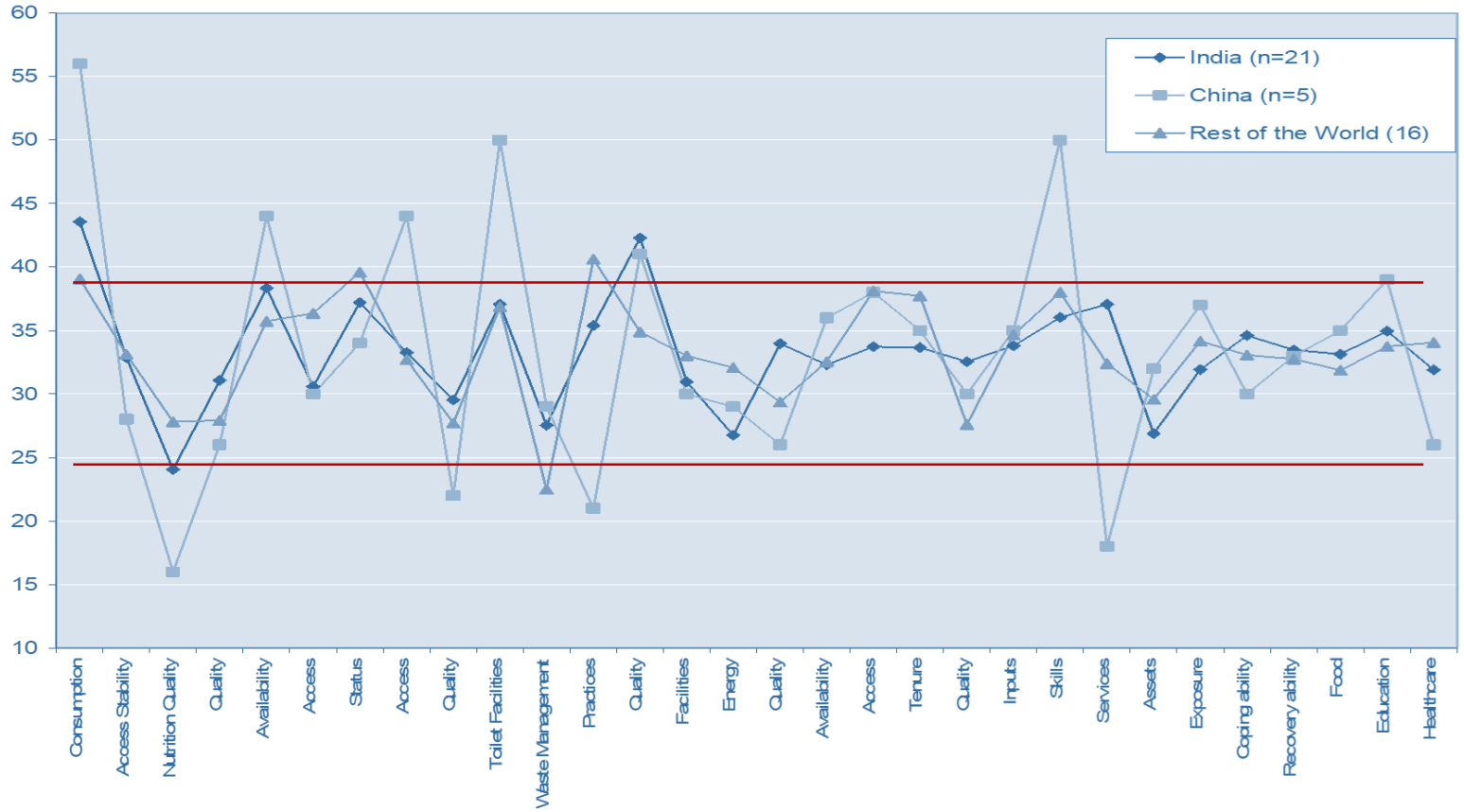
- Different distributions of weights across sub-components
- Expert opinions may significantly differ
- **Example**: Indian and Chinese experts compared to the rest of the world

BAP - Diversity of expert judgements



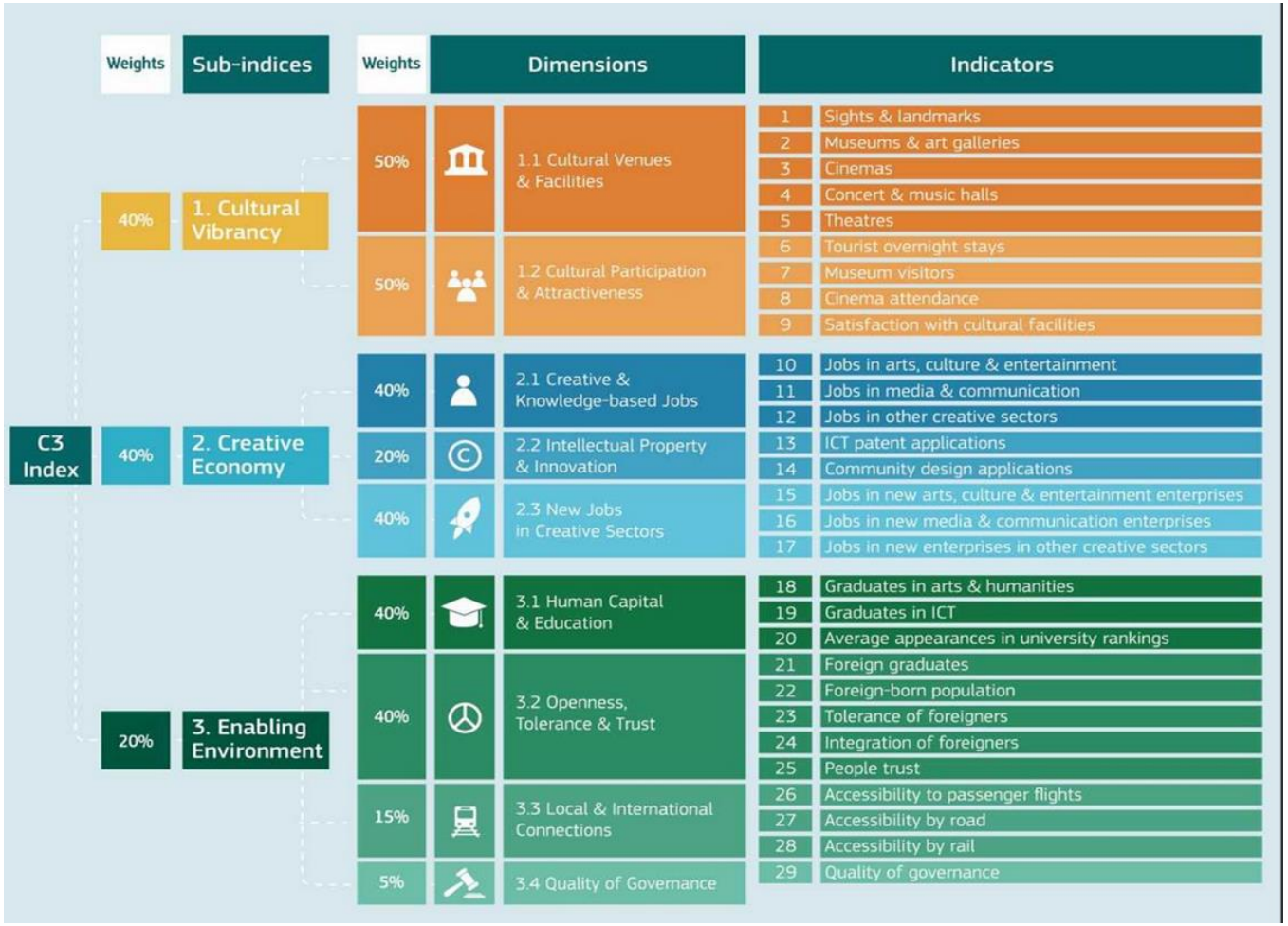
- Convergence: Indian experts have very similar views compared to the experts coming from the Rest of the world
- But …

BAP - Diversity of expert judgements



- Divergence: for 10 subcomponents Chinese experts have very different views
- Caution in averaging the judgements from different experts

BAP - Example 2: The Cultural and Creative Cities Monitor



European Commission

BAP - Example 2: The Cultural and Creative Cities Monitor

- **Selection of experts**: 17 Experts 5 from European Commission, 6 from Academia, 6 from international organisations

- Experts divided in 3 groups - weights assigned to the three sub-indices by each group:

Sub-index	Group 1	Group 2	Group 3	Average	Weight
Cultural vibrancy	40	50	40	43.33	**40**
Creative Economy	35	30	40	35.00	**40**
Enabling Environment	25	20	20	21.7	**20**

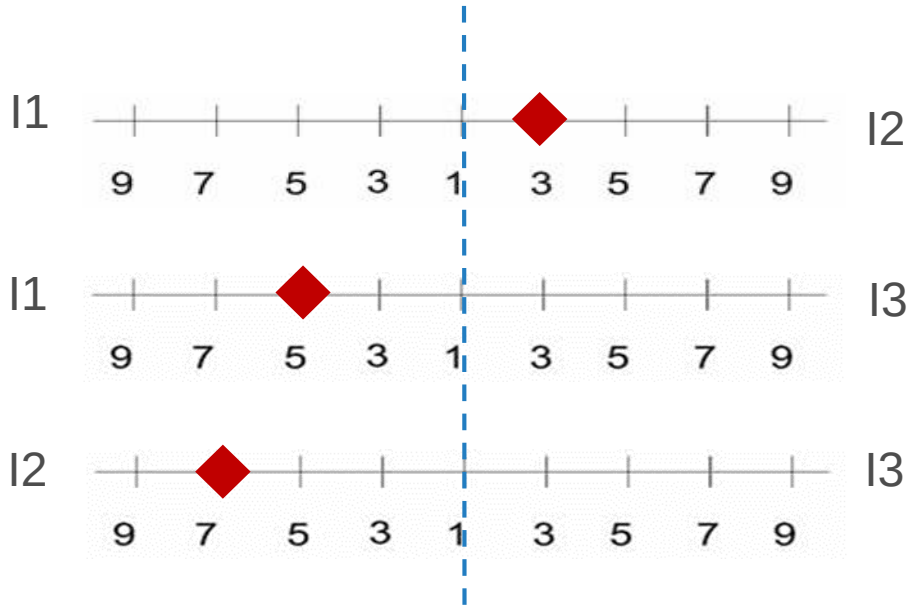
The experts agreed that that accessibility and governance dimensions (within the Enabling Environment sub-index) should have a minimum weight

Enabling Environment	Weight
Human capital & education	**40**
Openness, tolerance, trust	**40**
Connections	**15**
Quality of governance	**5**

Analytic Hierarchy Process (AHP)

- Multi-attribute decision making (MADM) method developed by Saaty (1987)

STEP 1. Question: *"Which of the two is more important?"* and *"by how much?"* on a scale ranging from **1** (**equally important**) to **9** (**most important):**



I2 results **3 times** more important than I1 (**I21 = 3**)

I1 results **5 times** more important than I3 (**I13 = 5**)

I2 results **7 times** more important than I3 (**I23 = 7**)

Analytic Hierarchy Process (AHP)

STEP 2. Set up a *N x N* matrix (M) with *N* being the number of indicators (in our case N = 3)

M	I1	I2	I3
I1	*1*	**1/3**	**5**
I2	**3**	*1*	**7**
I3	**1/5**	**1/7**	*1*

Elicited and **reciprocal** relative values (importance)

- The relative importance assigned to each pair of indicators has to be transformed in weights

STEP 3. Calculation of weights

- Normalized column method and geometric mean method (optimal for a limited number of indicators)
- Eigenvector technique

Normalized columns method

A. sum each column

M	I1	I2	I3
I1	1	1/3	5
I2	3	1	7
I3	1/5	1/7	1
sum	**21/5**	**31/21**	**13**

B. normalized rel. weights

M	I1	I2	I3
I1	5/21	7/31	5/13
I2	15/21	21/31	7/13
I3	1/21	3/31	1/13
sum	**1**	**1**	**1**

C. row average

Weights	
I1	**0.2828**
I2	**0.6434**
I3	**0.0738**

Geometric mean method

A. original elicited

M	I1	I2	I3
I1	1	1/3	5
I2	3	1	7
I3	1/5	1/7	1

B. geometric mean (GM)

GM	
I1	1.19
I2	2.76
I3	0.31
sum	**4.25**

C. normalized GM

Weights	
I1	**0.28**
I2	**0.6494**
I3	**0.0729**

Analytic Hierarchy Process (AHP)

- **Criticism of AHP**
  - Inconsistency forced by 1 to 9 scale (example: if *I12 = 4* and *I23 = 4*, then I13 should be at least *16*),
  - When the number of indicators is very large, AHP exerts cognitive stress on decision makers,
  - New alternatives can reverse the rank of existing alternatives.

- **Advantages of AHP**
  - Simplicity of pair-wise comparisons,
- Compared to BAP, the AHP method is less prone to errors of judgement => there is a possibility to check the consistency (in each set) of judgements and for each expert.

Analytic Hierarchy Process (AHP)

STEP 4. Checking the consistency

- **Consistency ratio (CR)**
  - Some inconsistency tolerated: CR ≤ 0.1 (10%) even though 0.2 may work as well (does not affect the weights drastically, Saaty, 1980),
  - Code in R available.

- **Alternative**: expert-based weights can be derived from **conjoint analysis (CA)**
  - frequently used in marketing and consumer research,
  - decompositional multivariate data analysis.

Thank you

matija.KOVACIC@ec.europa.eu | jrc-coin@ec.europa.eu

composite-indicators.jrc.ec.europa.eu

European Commission