A Multi-Criteria-Based Index for the Knowledge Economy in the EU25

Outline
• General aggregation issues
• Additive aggregation
• Geometric aggregation
• MCA
• The KBE Index for the EU15 + JP + US
• When to use what

Michaela Saisana
European Commission
Joint Research Centre of Ispra, Italy
michaela.saisana@jrc.it

Fourth Workshop on Indicators in the Knowledge Economy (KEI), Leuven, September 5-6, 2006
Based on:


Step 6 in the Handbook: (Weighting and) aggregation

Aggregation rules:

- Linear aggregation
- Geometric mean
- Multi-criteria analysis
Additive aggregation

### Summation of ranks
- The simplest method
- Based on ordinal information & independent to outliers **BUT** loses the absolute value information.

\[ CI_c = \sum_{q=1}^{Q} Rank_{qc} \]

### Number of indicators that are above and below some benchmark
- Uses nominal scores
- Threshold value \( p \) arbitrarily chosen
- Simple & unaffected by outliers **BUT** loses interval level information.

\[ CI_c = \sum_{q=1}^{Q} sgn \left[ \frac{I_{qc}}{I_{EUq}} - (1 + p) \right] \]

### Summation of weighted and normalized indicators
- By far the most widespread method
- Entails restrictions on the nature of indicators & weights
- Implies full (and constant) compensability
- Rewards indicators proportionally to the weights
- Requires normalisation
- Weights are trade offs not importance coefficients

\[ CI_c = \sum_{q=1}^{Q} w_q I_{qc} \]

\[ \sum_{q} w_q = 1 \text{ and } 0 \leq w_q \leq 1 \]
Additive aggregation

Example: Human Poverty Index 2001

\[ \text{HPI} = \left[ \frac{1}{3} \left( P_1^a + P_2^a + P_3^a \right) \right]^{1/a} ; a = 3 \]

- \( P_1 = \) Probability at birth of not surviving to age 40
- \( P_2 = \) Adult illiteracy rate
- \( P_3 = \) Unweighted average of population without sustainable access to an improved water source and children under weight for age

The ‘cubing’ i.e. \( a=3 \) ensures greater weight for the component with acute deprivation
3 dimension indices calculated for males and females and combined, penalizing differences in achievement

Equally distributed index =

\[
\left\{ \frac{\text{female popn. share} \times (\text{female index}^{1-\varepsilon})}{\text{male popn. share} \times (\text{male index}^{1-\varepsilon})} \right\}^{1/1-\varepsilon}
\]

where \( \varepsilon = 2 \) (moderate penalty for gender inequality)
Restrictions and assumptions

- Indicators need to be **mutually preferentially independent** (i.e. every subset of these indicators is preferentially independent of its complementary set of indicators) → very strong condition from both the **operational** and **epistemological** points of view.

- **Compensability** among the indicators is always assumed → complete substitutability among the various indicators
  
  E.g. in a sustainability index, economic growth can always substitute any environmental destruction or inside e.g., the environmental dimension, clean air can compensate for a loss of potable water. **From a descriptive point of view, such a complete compensability is often not desirable**

- **Weights** have the meaning of trade-off ratio. Yet, in all constructions of a composite indicator, weights are used as importance coefficients, as a consequence, **a theoretical inconsistency exists**.

- **Synergy or conflict** - Preferential independence implies that the trade-off ratio between two indicators is independent of the values of the n-2 other indicators
Example

A hypothetical composite: inequality, environmental degradation, GDP per capita and unemployment

Country A: 21, 1, 1, 1 \rightarrow 6
Country B: 6, 6, 6, 6 \rightarrow 6

Obviously the two countries would represent very different social conditions that would not be reflected in the composite.
Geometric aggregation

Example

A hypothetical composite: inequality, environmental degradation, GDP per capita and unemployment

\[ CI_c = \prod_{q=1}^{Q} x_q^{w_q} \]

Country A: 21, 1, 1, 1 → 2.14
Country B: 6, 6, 6, 6 → 6

- Countries with low scores in some indicators would prefer a linear rather than a geometric aggregation (the simple example above shows why).
- Yet, the marginal utility from an increase in low absolute score would be much higher than in a high absolute score under geometric aggregation

Country A: 21, 2, 1, 1 → 2.54 → 19% increase in the score
Country B: 6, 7, 6, 6 → 6.23 → 4% increase in the score

**Lesson:** a country should be more interested in increasing those sectors/activities/alternatives with the lowest score in order to have the highest chance to improve its position in the ranking if the aggregation is geometric rather than linear (Zimmermann and Zysno, 1983).
The absence of synergy or conflict effects among the indicators & weights express trade-offs between indicators are necessary conditions to admit either linear or geometric aggregation.
When different goals are equally legitimate and important, then a non compensatory logic may be necessary.

Example: physical, social and economic figures must be aggregated. If the analyst decides that an increase in economic performance can not compensate a loss in social cohesion or a worsening in environmental sustainability, then neither the linear nor the geometric aggregation are suitable.

Instead, a non-compensatory *multi-criteria approach* will assure non compensability by formalizing the idea of finding a compromise between two or more legitimate goals.

+ does not reward outliers
+ different goals are equally legitimate and important
+ no normalisation is required

**BUT**
- computational cost when the number of countries is high
### Multi-criteria type of aggregation

<table>
<thead>
<tr>
<th>Country</th>
<th>Indic.</th>
<th>GDP</th>
<th>Unemp. Rate</th>
<th>Solid wastes</th>
<th>Income disp.</th>
<th>Crime rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>25,000</td>
<td>0.15</td>
<td>0.4</td>
<td>9.2</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>45,000</td>
<td>0.10</td>
<td>0.7</td>
<td>13.2</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>20,000</td>
<td>0.08</td>
<td>0.35</td>
<td>5.3</td>
<td>80</td>
<td></td>
</tr>
</tbody>
</table>

**Weights:**

- A: 0.165
- B: 0.165
- C: 0.333

#### Linear aggregation: CBA

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.666</td>
<td>0.333</td>
</tr>
<tr>
<td>0.333</td>
<td>0</td>
<td>0.333</td>
</tr>
<tr>
<td>0.666</td>
<td>0.666</td>
<td>0</td>
</tr>
</tbody>
</table>

**Calculations:**

- ABC = 0.666 + 0.333 + 0.333 = 1.333
- BCA = 0.333 + 0.666 + 0.333 = 1.333
- CAB = 0.666 + 0.666 + 0.666 = 2
- ACB = 0.333 + 0.666 + 0.666 = 1.666
- BAC = 0.333 + 0.333 + 0.333 = 1
- CBA = 0.666 + 0.333 + 0.666 = 1.666

(Munda 2003, Munda & Nardo 2003)
The Computational problem

Moulin (1988, p. 312) clearly states that the Kemeny method is "the correct method" for ranking alternatives, and that the "only drawback of this aggregation method is the difficulty in computing it when the number of candidates grows".

With only 10 countries → 10! = 3,628,800 permutations
A NP-hard problem

The complexity class of decision problems that are intrinsically harder than those that can be solved by a nondeterministic Turing machine in polynomial time. When a decision version of a combinatorial optimization problem is proved to belong to the class of NP-complete problems, then the optimization version is NP-hard.

(definition given by the National Institute of Standards and Technology, http://www.nist.gov/dads/HTML/nphard.html)
This NP-hardness has discouraged the development of algorithms searching for exact solutions, thus the majority of the algorithms which have been proposed in the literature; are mainly

- **heuristics based on artificial intelligence**,  
- **branch and bound approaches and**  
- **multi-stage techniques**

(see e.g., Barthelemy et al., 1989; Charon et al., 1997; Cohen et al., 1999; Davenport and Kalagnam, 2004; Dwork et al., 2001; Truchon, 1998b).
A *new numerical algorithm* aimed at solving the computational problem connected to linear median orders by finding *exact solutions* has been proposed by Munda (2005).

- Linear median orders are computed by using their *theoretical equivalence with maximum likelihood rankings*.
- *Outranking matrices* are used as a starting computational step.
The Knowledge Economy Dataset (a continuously updated dataset...)

- **Series A** gives indicators for the four/five main drivers (see WP 1.1)
- **Series B** gives indicators for two types of outcomes: economic and social
- **Series C** gives additional indicators that could be useful in the scenario analyses.
A Component

27 Indicators

A1. Production and diffusion of ICT
- Investment in ICT (A1a4)
- # patent appl. to the EPO (A1a5)
- Broadband penetr. rate (A1c4)

A2. Human resources, skills and creativity:
- Pisa mathematical literacy of 15 year olds (A2a1)
- New PhDs per thousand population aged 25-34 (A2a4)
- Participation in LLL per working age population (25-64) (A2c3)
- Job to job mobility by NACE (A2e2)

A3. Knowledge production and diffusion:
- GERD expenditure/GDP (A3a1)
- GERD per capita (A3a3)
- Estimated Civil GERD as % of GDP (A3a4)
- GOVERD (I) (calculated) and HERD (II) as % of GDP (A3a5 I, II)
- GBAORD as % of GDP (calculated) (A3a6)
- BERD as a percentage of GDP (A3a7)
- BERD as a percentage of value added in industry (A3a9)
- BERD/GERD (calculated) (A3a10)
- Triadic patent families by priority year (A3b5 I, II)
- Share of all firms reporting public research (universities & institutes) as a high value information source (A3e1)
- Share of all firms reporting public research (universities & institutes) as a cooperation partner (A3e2)
- High tech exports/total exports (A3f2)

A4. Innovation, Entrepreneurship and creative destruction:
- Firm entries (A4a2)
- Firm exits (A4a3)
- Share of firms introducing new-to-market products (A3d1)
- Share of total sales from new-to-market products (A3d2)
- Share of total sales from new-to-firm products (A3d4)
- (A5a1)
Overall performance considers an equal importance to each component aggregated into 27 indicators.
Comparison of aggregation methods

- aggregation method affects principally the middle-of-the-road countries
- both aggregation schemes produce comparable rankings
- when compensability is not allowed, countries performing very poorly on some indicators, such as Finland or France see their rank lowered with respect to the linear aggregation, whereas countries that have less extreme values improve their situation, such as Belgium

\[ y = 0.9412x + 0.5294 \]

\[ R^2 = 0.8858 \]
A1. Production and diffusion of ICT

A2. Human resources, skills and creativity:

A3. Knowledge production and diffusion:

A4. Innovation, Entrepreneurship and creative destruction:
A1. Production and diffusion of ICT

A2. Human resources, skills and creativity:

A3 Knowledge production and diffusion:

A4. Innovation, entrepreneurship and creative destruction:
B Component

B1. Economic outputs:

- GDP per capita in PPS (B1a1)
- Real GDP growth rate (B1a2)
- Labour productivity per hour worked (B1b1)
- Gross fixed capital formation as % of GDP (calculated) (B1b3 I, II)
- Total employment growth (B1c1)

B2. Social performance:

- Energy intensity of the economy (B2a2)
- Employment rate of older workers (B2b1)
- Total employment rate (B1c2)
- Long term unemployment rate (B2b2)
- Inequality of income distribution (B2b4)
## Ranks - different components A, B

<table>
<thead>
<tr>
<th>Country</th>
<th>A</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>B1</th>
<th>B2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>13</td>
<td>11</td>
<td>13</td>
<td>12</td>
<td>9</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Belgium</td>
<td>11</td>
<td>4</td>
<td>10</td>
<td>13</td>
<td>12</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>Denmark</td>
<td>16</td>
<td>16</td>
<td>17</td>
<td>17</td>
<td>11</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>Germany</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>15</td>
<td>14</td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>Greece</td>
<td>17</td>
<td>17</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Spain</td>
<td>12</td>
<td>8</td>
<td>3</td>
<td>10</td>
<td>15</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Finland</td>
<td>10</td>
<td>1</td>
<td>14</td>
<td>1</td>
<td>13</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>France</td>
<td>14</td>
<td>9</td>
<td>11</td>
<td>9</td>
<td>16</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>Ireland</td>
<td>9</td>
<td>10</td>
<td>4</td>
<td>14</td>
<td>8</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>Italy</td>
<td>7</td>
<td>12</td>
<td>2</td>
<td>11</td>
<td>1</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>8</td>
<td>14</td>
<td>1</td>
<td>8</td>
<td>10</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>Netherlands</td>
<td>1</td>
<td>5</td>
<td>8</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Portugal</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Sweden</td>
<td>3</td>
<td>3</td>
<td>12</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>6</td>
<td>13</td>
<td>9</td>
<td>7</td>
<td>5</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>USA</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td>Japan</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>7</td>
<td>15</td>
<td>14</td>
</tr>
</tbody>
</table>
y = 0.5588x + 3.9706

$R^2 = 0.3123$
y = 0.402x + 5.3824

$R^2 = 0.1616$

Group B2

Netherlands
Portugal
Sweden
United Kingdom
Luxembourg
Spain
Belgium
Austria
Finland

Greece
Denmark
Germany

France

United Kingdom
Italy
Ireland

Finland

Portugal

USA
Japan

Belgium

Netherlands

Group B2
Analysis of different scenaria
when to use what?

When using a model or an algorithm to describe a real-world issue, formal coherence is a necessary property but not sufficient.

The model in fact should fit objectives and intentions of the user, i.e. it must be the most appropriate tool for expressing the set of objectives that motivated the whole exercise.

The choice of which indicators to use, how those are divided into classes, whether a normalization method has to be used (and which one), the choice of the weighting method, and how information is aggregated, all these features stem from a certain perspective on the issue to be modelled.
The absence of an “objective” way of constructing composites should not result in a rejection of whatever type of composite. Composites can meaningfully supply information provided that the relation between the framing of a problem and the outcome in the decision space are made clear.

A backward induction exercise could be useful in this context. Once the context and the modeller’s objectives have been made explicit, the user can verify whether and how the selected model fulfils those objectives.

No model is a priori better than another, provided internal coherence is assured. In practice, different models can meet different expectations and stakes. Therefore, stakes must be made clear, and transparency should guide the entire process.
Selected References

Books


Selected References

References – Recent papers


