The European Commission’s science and knowledge service

Joint Research Centre

Weights and Measures: Exploring and Optimising Composite Indicators with Sensitivity Analysis

William Becker

william.becker@jrc.ec.europa.eu
Composite indicators are aggregations of observable variables that aim to quantify non-measurable concepts.

**Why**

- Quantification of hard-to-quantify concepts
- Aggregation of a large number of indicators into an easily-digestible format
- Simple comparisons between countries/entities
Global Innovation Index (average)

Innovation Efficiency Ratio (ratio)

Innovation Input Sub-Index

Innovation Output Sub-Index

Y = \sum_{i=1}^{d} w_i x_{ji}, \quad j = 1, 2, \ldots, n

X_1 \quad X_2 \quad X_3 \quad X_4 \quad X_5 \quad X_6 \quad X_7
**Composite indicators**

- ✔ Simple* aggregations of indicators
- ✔ Very subjective
- ✔ Very uncertain
- ✔ Attract lots of public, political and media attention

*Actually not simple: statistical analysis, conceptual framework, robustness checks, etc...
\[ y_j = \sum_{i=1}^{d} w_i x_{ji}, \quad j = 1, 2, \ldots, n \]
“a general criticism that is frequently addressed at composite indicators, i.e. the arbitrary character of the procedures used to weight their various components”

- the Stiglitz Report, 2010
Two ways to weight

Weights by belief

- Equal weights
- Budget allocation
- Public opinion
- Analytic hierarchy process
- Conjoint analysis

Weights by Statistical models

- Principal component analysis
- Factor analysis
- Data envelopment analysis
- Regression approach
Weights are typically assigned to reflect importance, but...

weights do not equal measures
A composite indicator to rank teachers...

\[ y = \frac{1}{3} (x_1 + x_2 + x_3) \]

\( x_1 = \) number of publications
\( x_2 = \) teaching feedback
\( x_3 = \) hours of teaching and office work

All have been standardised to have unit variance, but \( x_2 \) and \( x_3 \) have a correlation of 0.7. After calculating the values of the composite indicator, \( R^2 \) is used to check influence:

\[ R_i^2 := \text{corr}^2(y, x_i) \]

\[ R_1^2 = 0.227 \]
\[ R_2^2 = 0.657 \]
\[ R_3^2 = 0.657 \]

Increase the weight of \( x_1 \)? Teachers will complain that the index unfairly favours publications!

Correlations are very common in composite indicators.
How can we measure importance?
Correlation coefficient

\[ R_i := \text{corr}(y, x_i) := \frac{\text{cov}(y, x_i)}{\sigma_y \sigma_i}. \]

\[ R^2_i := \text{corr}^2(y, x_i), \]

Ok but only measures linear dependence. Not always the case.

Correlation ratio*

\[ S_i \equiv \eta_i^2 := \frac{X_i \left( \mathbb{E}_{x \sim i} (y \mid x_i) \right)}{V(y)}, \]

Also known as “main effect index”, “first-order sensitivity index”, “nonlinear R^2”...

- Allows for nonlinear dependence
- Easily estimated by regression

*First conceived by Karl Pearson in 1905
Nonlinearity in the main effect

Nonlinear regression approaches required
\[ S_i \equiv \eta_i^2 := \frac{\text{var}_{x_i} (E_{x \sim i} (y \mid x_i))}{\text{var}(y)}, \]

\[ \sum_{j=1}^{n} (\hat{y}_j - \bar{y}) \]

\[ \sum_{j=1}^{n} (y_j - \bar{y}) \]

\( \rho(x) \) unknown:
Arbitrarily-distributed dependent variables.
Nonlinear regression approaches
1. Penalised smoothing splines

\[ \beta_0 + \beta_1 x + \ldots + \beta_p x^p + \sum_{k=1}^{K} \beta_{pk} (x - \kappa_k)^p \]

- **Polynomial part**
- **Spline part**

The two components are weighted by a smoothing term.
1. Penalised smoothing splines

- Smooth fits
- Fast parameter estimation using linear regression properties: can deal with large datasets

\[ \beta_0 + \beta_1 x + \ldots + \beta_p x^p + \sum_{k=1}^{K} \beta_{pk} (x - \kappa_k)^p \]
2. Local linear regression

\[ \hat{m}(x_i) = \frac{\sum_{j=1}^{n} w(x_{ji} - x_i; h) y_j}{\sum_{j=1}^{n} w(x_{ji} - x_i; h)} \]

3. Gaussian process regression (kriging)

A probabilistic distribution over functions.

Points are assumed to be distributed joint-normally.

Bayesian approach to fitting: sample from posterior using MCMC.

Posterior draws can be used to build a distribution over $S_j$.

$$f(x)|\sigma^2, b, \sigma_n^2 \sim \mathcal{GP}(m(x), \sigma^2 c(x, x') + \sigma_n^2)$$
Can we do more?
Separating correlation from aggregation

\[ S_i = S^u_i + S^c_i \]

**Correlation ratio**  Uncorrelated part  Correlated part

- Reveals variables that could be removed (uncorrelated contribution is low)
- Reveal indicators which are not sensitive to changes in weights
Separating correlation from aggregation

We regress $x_i$ on to the remaining variables $x_{\sim i}$.
Separating correlation from aggregation

We regress $x_i$ on to the remaining variables $x_{\sim i}$

Then subtract from $x_i$: this effectively removes the dependence on other variables.

\[ S^c_i = S_i - S^u_i \]

• Now regress $y$ on $\hat{z}_i$ (using NL regression) to get uncorrelated part.

\[ \hat{z}_i = x_i - \hat{x}_i = x_i - \beta_0 + \sum_{l \neq i} \hat{\beta}_l x_l \]
Separating correlation from aggregation

\[ \hat{z}_i = x_i - \hat{x}_i = x_i - \beta_0 + \sum_{l \neq i} \hat{\beta}_l x_l \]

\[ S_c^i = S_i - S_u^i \]

\[ \frac{\sum_{j=1}^{n} (\hat{y}_j - \bar{y})}{\sum_{j=1}^{n} (y_j - \bar{y})} \]

- Now regress \( y \) on \( \hat{z}_i \) (using NL regression) to get uncorrelated part.
- Can also use nonlinear regression to remove nonlinear dependence
Can we tune the weights?
\[ w_{\text{opt}} = \arg\min_w \sum_{i=1}^{d} (\tilde{S}_i^* - \tilde{S}_i(w))^2. \]
Summary

• **Estimate** $S_i$ **of** indicators using nonlinear regression (allows for correlation)
• **Separate correlated part** of $S_i$ **using** a regression approach (generalised to nonlinear dependence using GPs)
• **Optimise weights** to agree with target “importance” using numerical algorithm
Back to the beginning
<table>
<thead>
<tr>
<th>Academic Reputation</th>
<th>Employer Reputation</th>
<th>Faculty/Student Ratio</th>
<th>Citations per Faculty</th>
<th>International Faculty</th>
<th>International Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>40%</td>
<td>10%</td>
<td>20%</td>
<td>20%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>Opinion-based survey</td>
<td>Opinion-based survey</td>
<td>Teacher/student ratio</td>
<td>Citations divided by faculty</td>
<td>Proportion of international faculty members</td>
<td>Proportion of international students</td>
</tr>
</tbody>
</table>
All of the largest rank differences labelled here are such that the Times rank is higher than the QS rank.
Introduction  Weights  Measures  NL Regression  Correlation  Optimisation  University Rankings

QS Regression

x1  Academic Reputation
x2  Employer Reputation
x3  Faculty Student
x4  Citations per Faculty
x5  International Faculty
x6  International Students
QS Regression

x1 Academic Reputation
x2 Employer Reputation
x3 Faculty Student
x4 Citations per Faculty
x5 International Faculty
x6 International Students
Times Regression

x1 Citations
x2 Industry Income
x3 International Outlook
x4 Research
x5 Teaching
Times Regression

- $x_1$: Citations
- $x_2$: Industry Income
- $x_3$: International Outlook
- $x_4$: Research
- $x_5$: Teaching
Times target importance compared to Si

- **Target**
- **Original**

Citations | Industry Income | International Outlook | Research | Teaching

---

Introduction  | Weights  | Measures  | NL Regression  | Correlation  | Optimisation  | University Rankings

European Commission
### Times target importance compared to Si

- **Target**
- **Original**
- **Optimised (any weights)**
- **Optimised (pos weights)**

<table>
<thead>
<tr>
<th></th>
<th>Citations</th>
<th>Industry Income</th>
<th>International Outlook</th>
<th>Research</th>
<th>Teaching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>0.30</td>
<td>0.03</td>
<td>0.08</td>
<td>0.30</td>
<td>0.30</td>
</tr>
<tr>
<td>Optimised</td>
<td>0.46</td>
<td>-0.08</td>
<td>0.19</td>
<td>0.03</td>
<td>0.39</td>
</tr>
<tr>
<td>Opt (+)</td>
<td>0.44</td>
<td>0.11</td>
<td>0.16</td>
<td>0.00</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>S_i</td>
<td></td>
<td>S_i,u</td>
<td></td>
<td>S_i,c</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>-------</td>
<td>---------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>Linear</td>
<td>Spline</td>
<td>Loc. Lin.</td>
<td>GP</td>
<td>Linear</td>
</tr>
<tr>
<td>x1</td>
<td>Academic Reputation</td>
<td>0.67</td>
<td>0.75</td>
<td>0.78</td>
<td>0.75</td>
</tr>
<tr>
<td>x2</td>
<td>Employer Reputation</td>
<td>0.45</td>
<td>0.48</td>
<td>0.48</td>
<td>0.46</td>
</tr>
<tr>
<td>x3</td>
<td>Faculty Student</td>
<td>0.14</td>
<td>0.17</td>
<td>0.15</td>
<td>0.16</td>
</tr>
<tr>
<td>x4</td>
<td>Citations per Faculty</td>
<td>0.16</td>
<td>0.17</td>
<td>0.16</td>
<td>0.17</td>
</tr>
<tr>
<td>x5</td>
<td>International Faculty</td>
<td>0.06</td>
<td>0.10</td>
<td>0.09</td>
<td>0.08</td>
</tr>
<tr>
<td>x6</td>
<td>International Students</td>
<td>0.09</td>
<td>0.10</td>
<td>0.08</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Times</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x1</td>
<td>Citations</td>
<td>0.27</td>
<td>0.43</td>
<td>0.43</td>
<td>0.41</td>
</tr>
<tr>
<td>x2</td>
<td>Industry Income</td>
<td>0.02</td>
<td>0.05</td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td>x3</td>
<td>International Outlook</td>
<td>0.01</td>
<td>0.03</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td>x4</td>
<td>Research</td>
<td>0.88</td>
<td>0.90</td>
<td>0.90</td>
<td>0.90</td>
</tr>
<tr>
<td>x5</td>
<td>Teaching</td>
<td>0.85</td>
<td>0.86</td>
<td>0.86</td>
<td>0.85</td>
</tr>
</tbody>
</table>

- NL regression estimates are fairly similar but linear regression not always sufficient
- Uncorrelated part can be dominated by correlated part: more likely that variable is not contributing as intended and difficult to optimise
QS with optimised weights
<table>
<thead>
<tr>
<th>New Rank</th>
<th>Original Rank</th>
<th>University</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>California Institute of Technology (Caltech)</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>Massachusetts Institute of Technology (MIT)</td>
</tr>
<tr>
<td>3</td>
<td>17</td>
<td>Johns Hopkins University</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>University of Cambridge</td>
</tr>
<tr>
<td>5</td>
<td>66</td>
<td>University of Illinois at Urbana-Champaign</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>University of Oxford</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>Stanford University</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>UCL (University College London)</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>Harvard University</td>
</tr>
<tr>
<td>10</td>
<td>24</td>
<td>Duke University</td>
</tr>
</tbody>
</table>

**QS with optimised weights**
Introduction  Weights  Measures  NL Regression  Correlation  Optimisation  University Rankings

Original rank

Rank with (unconstrained) optimised weights

Times with optimised weights
### New Rank | Original Rank | University
--- | --- | ---
1 | 5 | Massachusetts Institute of Technology
2 | 1 | University of Oxford
3 | 2 | California Institute of Technology
4 | 8 | Imperial College London
5 | 3 | Stanford University
6 | 4 | University of Cambridge
7 | 17 | Johns Hopkins University
8 | 7 | Princeton University
9 | 9 | ETH Zurich
10 | 18 | Duke University

**Times with optimised weights**
Conclusions

Weights don’t equal importance in composite indicators.

Sensitivity analysis can help to explore and improve rankings

Rankings very sensitive to assumptions. Many other assumptions to explore here.

Open Questions

Does importance = $S_i$?

When we assign “importances” to variables, are we implicitly taking some correlation into account?
References

**Composite indicators and Sensitivity Analysis**


**Sensitivity indices using nonlinear regression**


**Separating correlations using regression**


**Bayesian Gaussian processes**


**Splines**


**Local polynomial regression**