The European Commission’s science and knowledge service

Joint Research Centre

European Commission

Step 8: Quality Assurance and Robustness

William Becker

COIN 2017 - 15th JRC Annual Training on Composite Indicators & Scoreboards
06-08/11/2017, Ispra (IT)


rankings are here to stay, and it is therefore worth the time and effort to get them right

Some feedback

Rankings range from irresponsible musings by self-appointed experts and money-making schemes by commercial organizations to, at their best, serious efforts by academic or research organizations (Aitbach, 2015)

Notwithstanding recent attempts to establish good practice in composite indicator construction (OECD, 2008), "there is <u>no recipe for building composite indicators</u> that is at the same time <u>universally applicable and sufficiently detailed</u>" (Cherchye et al., 2007).

Booysen (2002, p.131) summarises the debate on composite indicators by noting that "<u>not one single element of the methodology of composite indexing is above criticism</u>".

Andrews et al. (2004)] argue that "many indices rarely have adequate scientific foundations to support precise rankings: [...] typical practice is to acknowledge uncertainty in the text of the report and then to present a table with unambiguous rankings"

Tools for “serious efforts”

Quality Assurance

Robustness

Tools for “serious efforts”

Quality Assurance

Robustness

Tools for “serious efforts”

SELF DEFENSE

Quality Assurance

Ensuring statistical coherence (see previous lectures)

How do indicators contribute to the composite?

Robustness

How sensitive is the composite indicator to its assumptions?

- Transparency
- Exploration of uncertainty
- Anticipate criticism

We use Monte Carlo methods and sensitivity analysis

European Commission

THE TIMES HIGHER EDUCATION WORLD UNIVERSITY RANKINGS

tax justice network

financial secrecy index

THE FREEDOM INDEX

The Good Country Index

GPI GLOBAL PEACE INDEX

OECD Better Life Index

Natural Resource Governance Institute

A Measure of Transparency and Accountability in the Oil, Gas and Mining Sector

INSEAD The Business School for the World

human capital leadership institute

Adecco better work, better life

The Global Talent Competitiveness Index

Growing talent for today and tomorrow

Environmental Performance Index

QS WORLD UNIVERSITY RANKINGS

THE HUMAN FREEDOM INDEX

A GLOBAL MEASUREMENT OF PERSONAL, CIVIL, AND ECONOMIC FREEDOM

THE GLOBAL SLAVERY INDEX 2016

THE GLOBAL INNOVATION INDEX

TIMES HIGHER EDUCATION
WORLD UNIVERSITY RANKINGS

tax justice network

financial secrecy index

slide added by PowerPoint Labs

THE FREEDOM INDEX

The Good Country Index

GPI
GLOBAL PEACE INDEX

OECD Better Life Index

Natural Resource Governance Institute

A Measure of Transparency and Accountability in the Oil, Gas and Mining Sector

INSEAD The Business School for the World
human capital leadership institute
Adecco better work, better life
The Global Talent Competitiveness Index
Growing talent for today and tomorrow

Environmental Performance Index

QS WORLD UNIVERSITY RANKINGS®

THE HUMAN FREEDOM INDEX
A GLOBAL MEASUREMENT OF PERSONAL, CIVIL, AND ECONOMIC FREEDOM

THE GLOBAL SLAVERY INDEX 2016

THE GLOBAL INNOVATION INDEX

TIMES HIGHER EDUCATION
WORLD UNIVERSITY RANKINGS

QS WORLD UNIVERSITY RANKINGS®

European Commission

Definition of the university *is broad*

A university – as the name suggests – tends to encompass a broad range of purposes and dimensions, focus and missions. It is difficult to condense into a compact measure.

But measuring and benchmarking excellence is still in demand:

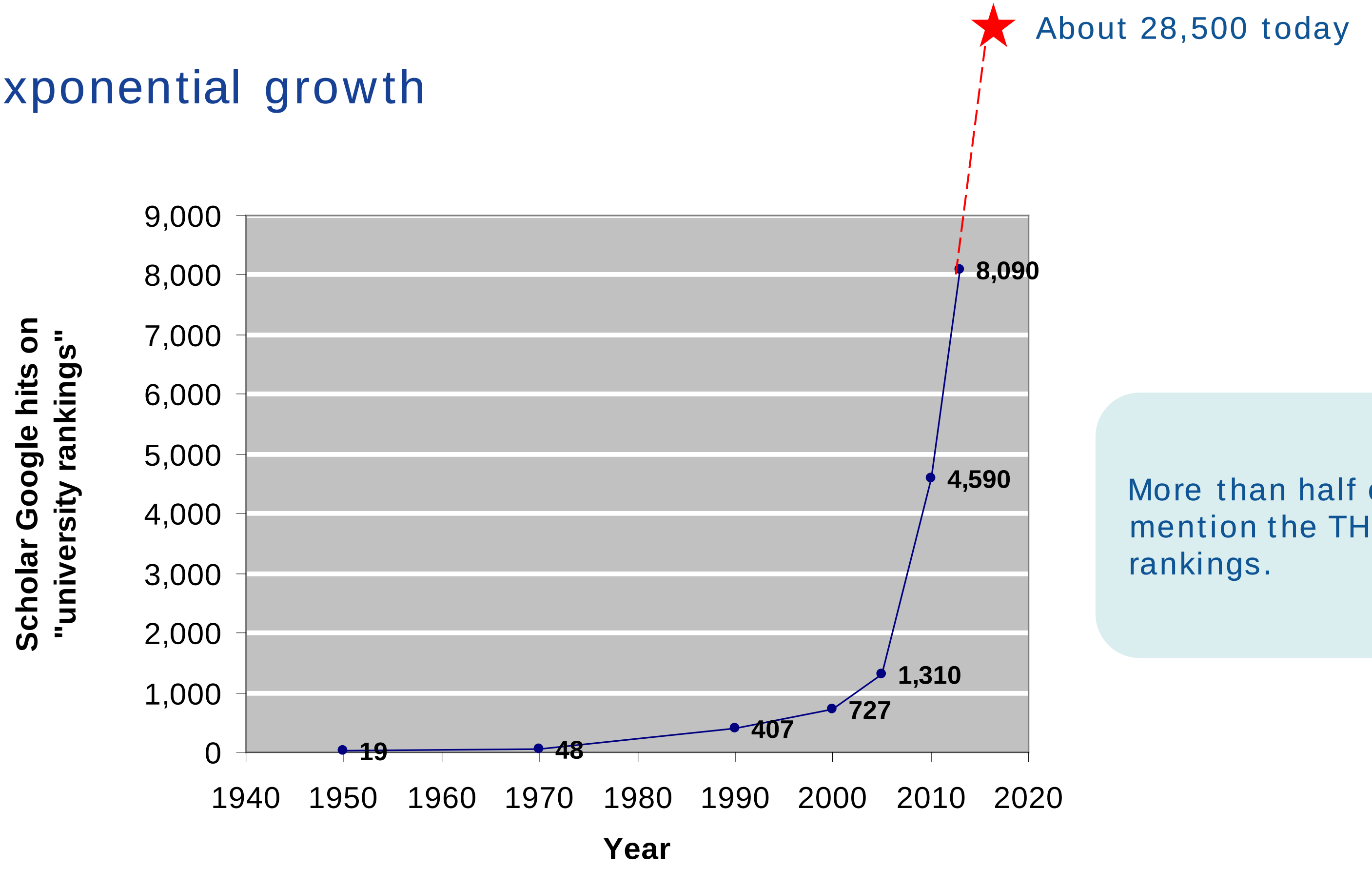
- Governance
- Accountability
- Transparency

The growing mobility of students and researchers has also created a market for these measures among the prospective students and their families.

Global rankings have raised debates and policy responses (EU, national level):

- To improve the positioning of a country within the existing measures
- To create new measures
- To discuss regional performance (e.g. show that USA is well ahead of Europe in terms of cutting-edge university research)

Exponential growth



More than half of these sources mention the THES and QS rankings.

Classement de Shangai : les universités françaises à la traîne

Seules 23 universités françaises figurent dans le classement de Shanghai des 500 meilleurs établissements mondiaux.

C'est un électrochoc qui secoue chaque année les universités tricolores en pleine torpeur estivale. Le classement de Shanghai, qui évalue les performances des meilleurs établissements d'enseignement supérieur mondiaux, vient d'être dévoilé. Ni chute ni progression spectaculaire : avec 23 universités dans le top 500 (22 l'an passé), et 3 dans le top 100 (4 l'an passé), la France se classe au septième rang des 37 pays, rétrocédant une place à la Suède. Des résultats décevants : en légère baisse par rapport à l'an passé, la France ne parvient toujours pas à rattraper son retard sur ses homologues britanniques et allemandes, dont une quarantaine d'universités sont classées.

Palmarès mondial des universités

Rang	Institution	Pays
1	Harvard	Etats-Unis
2	Stanford	Etats-Unis
3	Berkeley	Etats-Unis
4	Cambridge	Roy.-Uni
5	Massachusetts Inst. Tech. (MIT)	Etats-Unis
6	California Inst. Tech.	Etats-Unis
7	Columbia	Etats-Unis
8	Princeton	Etats-Unis
9	Chicago	Etats-Unis
10	Oxford	Roy.-Uni
42	Univ. Paris-VI	France
49	Univ. Paris-XI	France
73	Ecole normale sup. de Paris	France

Source : université de Jiao Tong.

Strasbourg-I ne peut qu'espérer que

Comment sauver l'université française

Des pistes pour remédier à la rétrogradation de nos établissements d'enseignement supérieur dans les classements internationaux

Jacques Blamont

Membre de l'Académie des sciences

Le classement de Shanghaï, qui fait désormais référence quels que soient ses défauts, vient de rétrograder encore une fois les universités françaises par rapport à leurs sœurs étrangères. Dans les cent premières, il n'en accepte plus que trois : Paris-VI (42e), Paris-XI (49e) et l'Ecole normale supérieure (73e).

0 ; revenus : 0. Ses ressources totales seraient donc sept fois inférieures à celles du Michigan. Les chiffres relatifs à d'autres universités américaines donne-

les entreprises abonderaient les bourses. La gestion des droits d'inscription et des bourses ressortirait à l'autonomie universitaire.

Une deuxième leçon à tirer de la pratique internationale est que l'université ne doit pas être asphyxiée par des gens qui n'y ont pas leur place. Une sélection à l'entrée s'impose, soit par une meilleure utilisation du baccalauréat (seuls seraient admis les titulaires d'une mention), soit

University rankings are used to judge the performance of university systems …

whether intended or not on by their proponents

European Commission

Classement de Shangai : les universités françaises à la traîne

Seules 23 universités françaises figurent dans le classement de Shanghai des 500 meilleurs établissements mondiaux.

C'est un électrochoc qui secoue chaque année les universités tricolores en pleine torpeur estivale. Le classement de Shanghai, qui évalue les performances des meilleurs établissements d'enseignement supérieur mondiaux, vient d'être dévoilé. Ni chute ni progression spectaculaire : avec 23 universités dans le top 500 (22 l'an passé), et 3 dans le top 100 (4 l'an passé), la France se classe au septième rang des 37 pays, rétrocédant une place à la Suède. Des résultats décevants : en légère baisse par rapport à l'an passé, la France ne parvient toujours pas à rattraper son retard sur ses homologues britanniques et allemandes, dont une quarantaine d'universités sont classées.

Palmarès mondial des universités

Rang	Institution	Pays
1	Harvard	Etats-Unis
2	Stanford	Etats-Unis
3	Berkeley	Etats-Unis
4	Cambridge	Roy.-Uni
5	Massachusetts Inst. Tech. (MIT)	Etats-Unis
6	California Inst. Tech.	Etats-Unis
7	Columbia	Etats-Unis
8	Princeton	Etats-Unis
9	Chicago	Etats-Unis
10	Oxford	Roy.-Uni
42	Univ. Paris-VI	France
49	Univ. Paris-XI	France
73	Ecole normale sup. de Paris	France

Source : université de Jiao Tong.

Strasbourg-I ne peut qu'espérer que

France:

- Creation of 10 centres of HE excellence
- The minister of Education set a target to put at least 10 French universities among the top 100 in ARWU by 2012
- President has put French standing in these international ranking at the forefront of the policy debate (Le Monde, 2008). -→ (2016: there are still three)

Italy (0 Uni in the top 100 of the ARWU ranking→ seen as failure of the national educational system).

Spain (1 Uni in the top 200 of the ARWU → hailed as a great national achievement)

UK universities tumble in world rankings amid Brexit concerns

Uncertainty over research funding and immigration rules blamed for decline, as Cambridge slips out of top three for first time

- Top 200 universities in the world - the table

Cambridge University, now ranked fourth in the world. Photograph: Bloomberg via Getty Images

> The surveys that informed the rankings were carried out before the UK voted in June to leave the EU

An extreme impact of global rankings

What - 2005 THES created a major controversy in Malaysia: country's top two universities slipping by almost 100 places compared to 2004.

Why - change in the ranking methodology (not well known fact and of limited comfort)

Impact - Royal commission of inquiry to investigate the matter. A few weeks later, the Vice-Chancellor of the University of Malaysia stepped down.

Composite indicators move money

(policy-makers are listening)

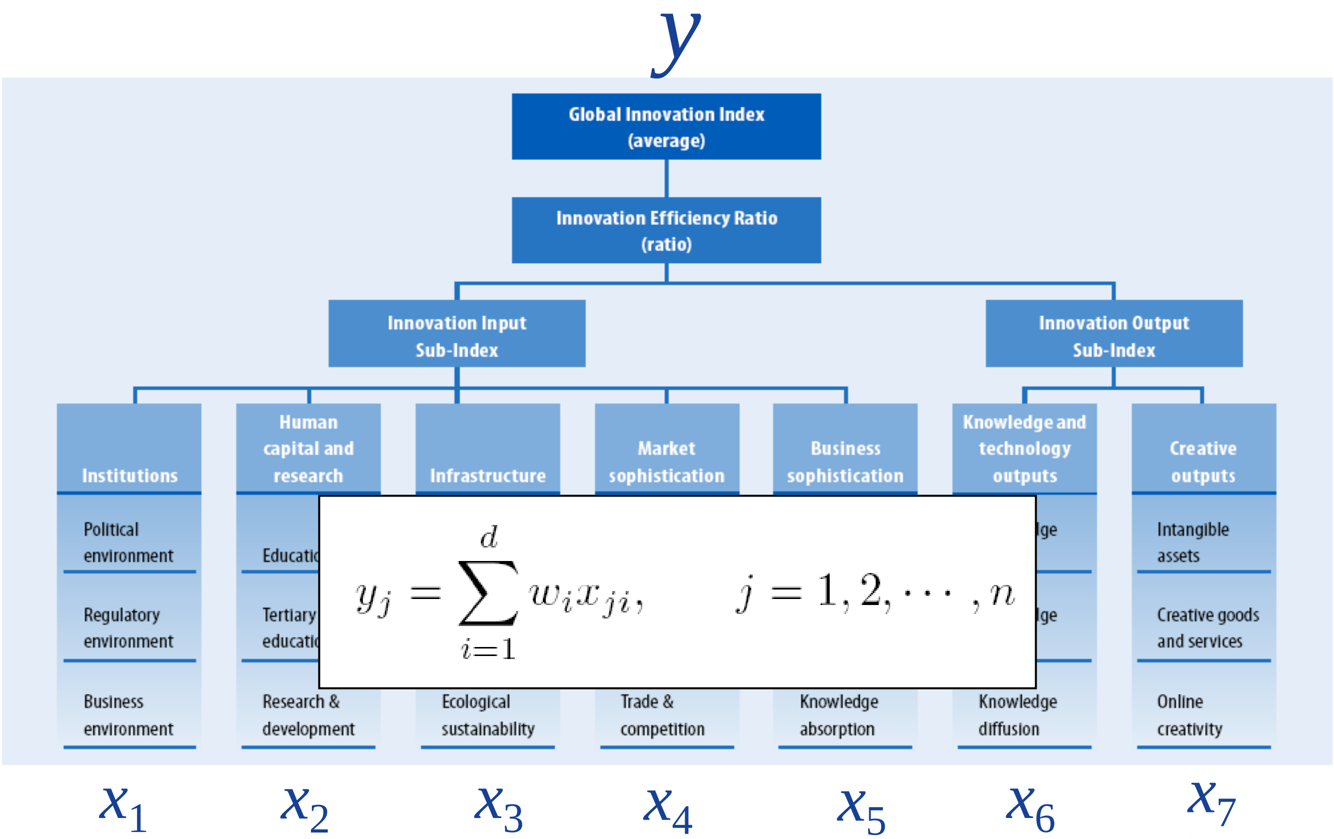
Composite indicator	Budget programme	2018 Budget
International Logistics Performance Index	Customs 2020	80.2 Million
Inform Index for Risk Management	Humanitarian aid	1.1 Billion
Corruption index, Press freedom Index, Ease of Doing Business Index + others.	Instrument for Pre-accession Assistance	1.7 Billion
Worldwide Governance Indicator	Development Cooperation Instrument	3.0 Billion

European Commission

Part 1: Quality assurance— Weighting

“a general criticism that is frequently addressed at composite indicators, i.e. <u>the arbitrary character of the procedures used to weight their various components</u>”

- the Stiglitz Report, 2010



$$y_j = \sum_{i=1}^{d} w_i x_{ji}, \qquad j = 1, 2, \cdots, n$$

$$\boxed{y_j = \sum_{i=1}^{d} w_i x_{ji}, \qquad j = 1, 2, \cdots, n}$$

Two ways to weight

Weights by belief

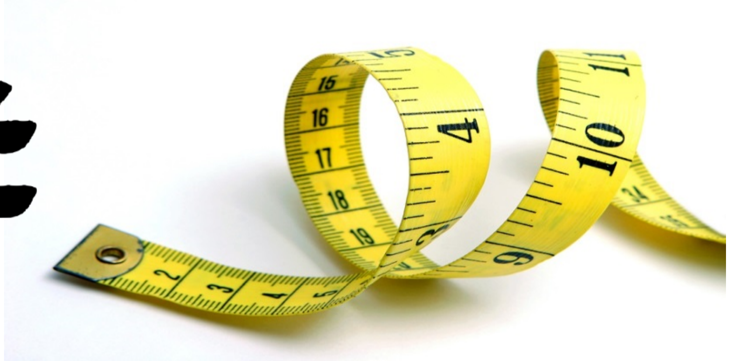
- Equal weights
- Budget allocation
- Public opinion
- Analytic hierarchy process
- Conjoint analysis

Weights by Statistical models

- Principal component analysis
- Factor analysis
- Data envelopment analysis
- Regression approach

Weights are typically assigned to reflect importance, but…

≠

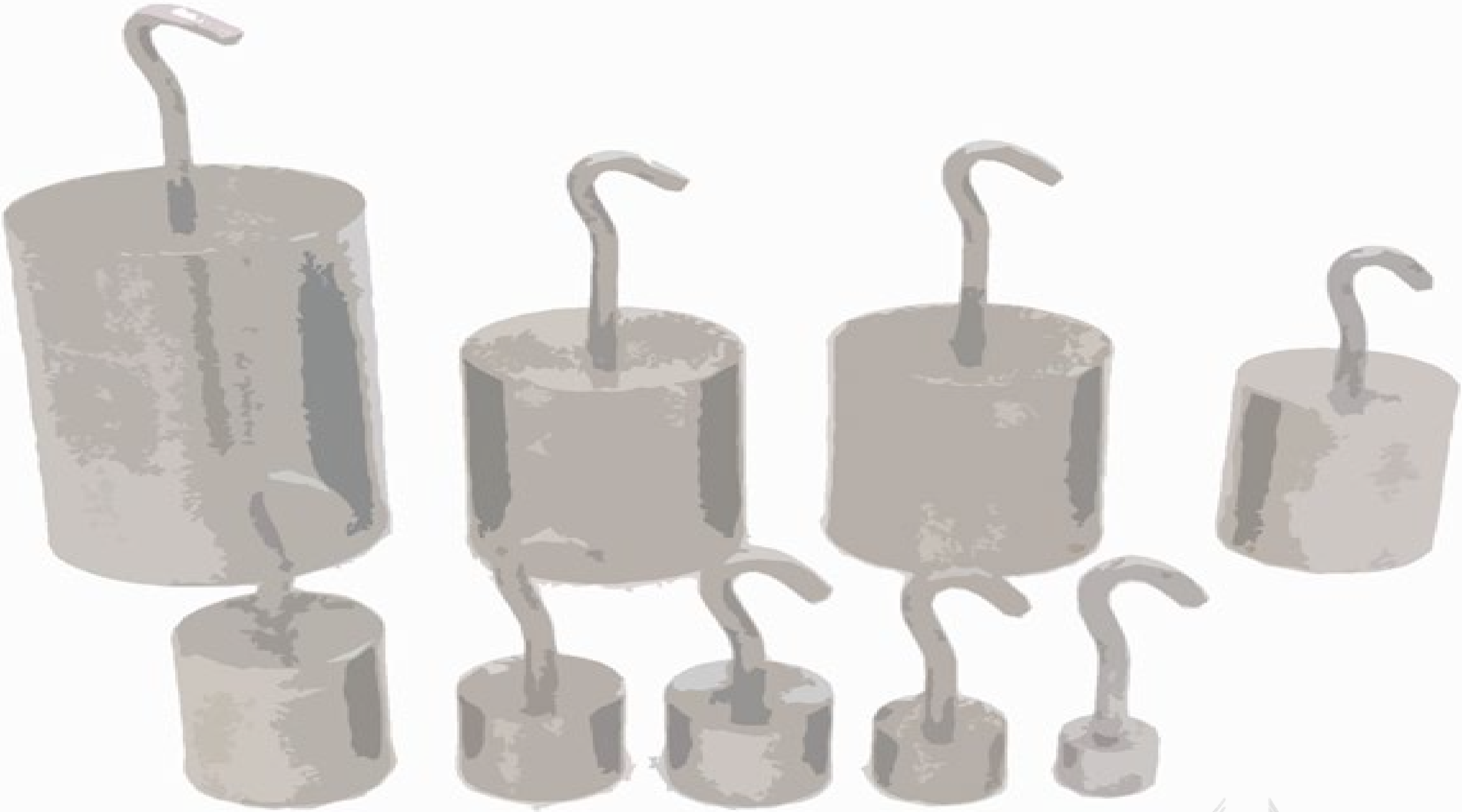


weights do not equal measures

Weights are typically assigned to reflect importance, but...

≠

weights do not equal measures



A composite indicator to rank teachers…

European Commission

A composite indicator to rank teachers…

A composite indicator to rank teachers…

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

$$\boxed{y = \frac{1}{3}(x_1 + x_2 + x_3)}$$

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:

$$\boxed{R_i^2 := \mathrm{corr}^2(y, x_i)} \longrightarrow$$

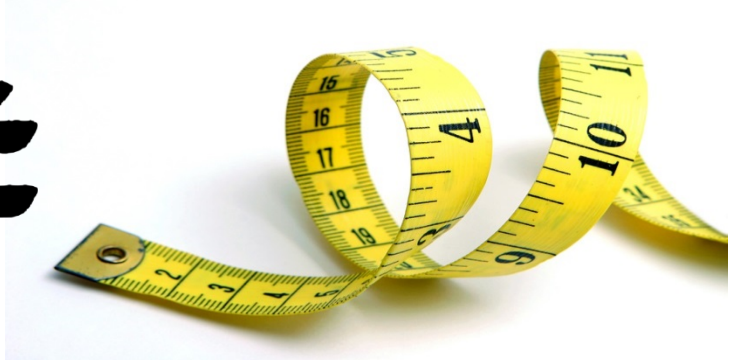
R_1^2=0.227
R_2^2=0.657
R_3^2=0.657

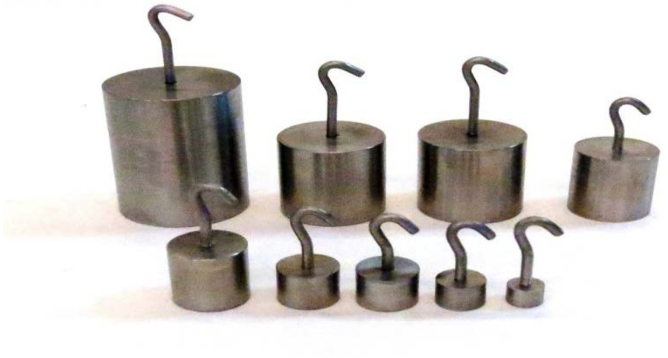
Increase the weight of x_1? Teachers will complain that the index unfairly favours publications!

Correlations are very common in composite indicators.

European Commission

≠





≠

How can we measure importance?

Measuring the importance of indicators

Correlation coefficient

$$R_i := \mathrm{corr}(y, x_i) := \frac{\mathrm{cov}(y, x_i)}{\sigma_y \sigma_i}.$$

$$R_i^2 := \mathrm{corr}^2(y, x_i),$$

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

Correlation ratio*

$$S_i \equiv \eta_i^2 := \frac{\mathrm{V}_{x_i}\left(\mathrm{E}_{\mathbf{x}_{\sim i}}\left(y \mid x_i\right)\right)}{\mathrm{V}\left(y\right)},$$

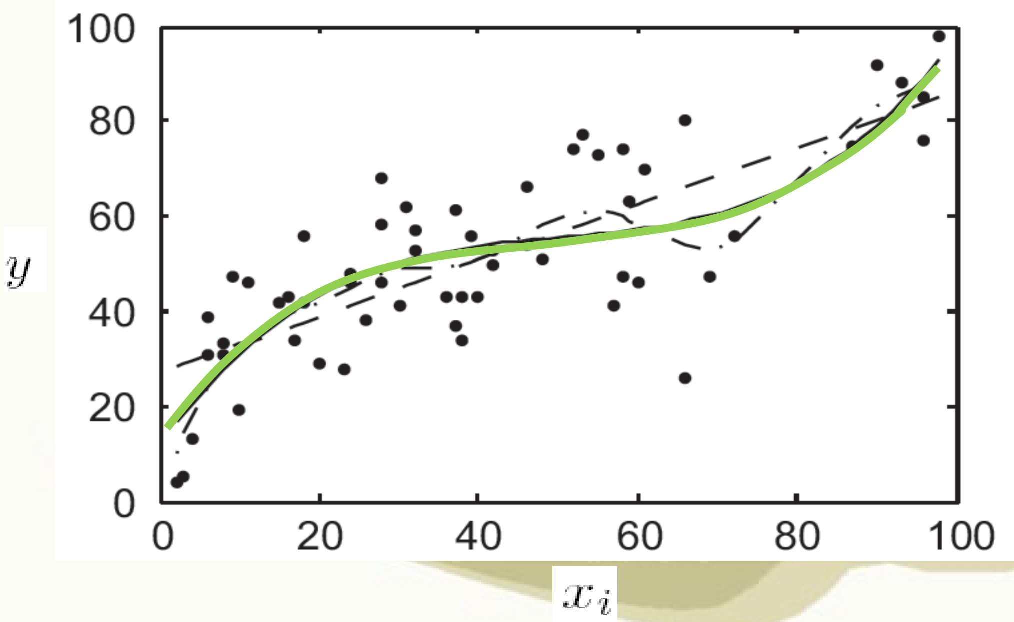
Also known as "main effect index", "first-order sensitivity index", "nonlinear R^2"…

- Allows for nonlinear dependence
- Easily estimated by regression

*First conceived by Karl Pearson in 1905

$$S_i \equiv \eta_i^2 := \frac{\boxed{V_{x_i}\left(E_{\mathbf{x}_{\sim i}}\left(y \mid x_i\right)\right)}}{V(y)},$$

Nonlinearity in the main effect

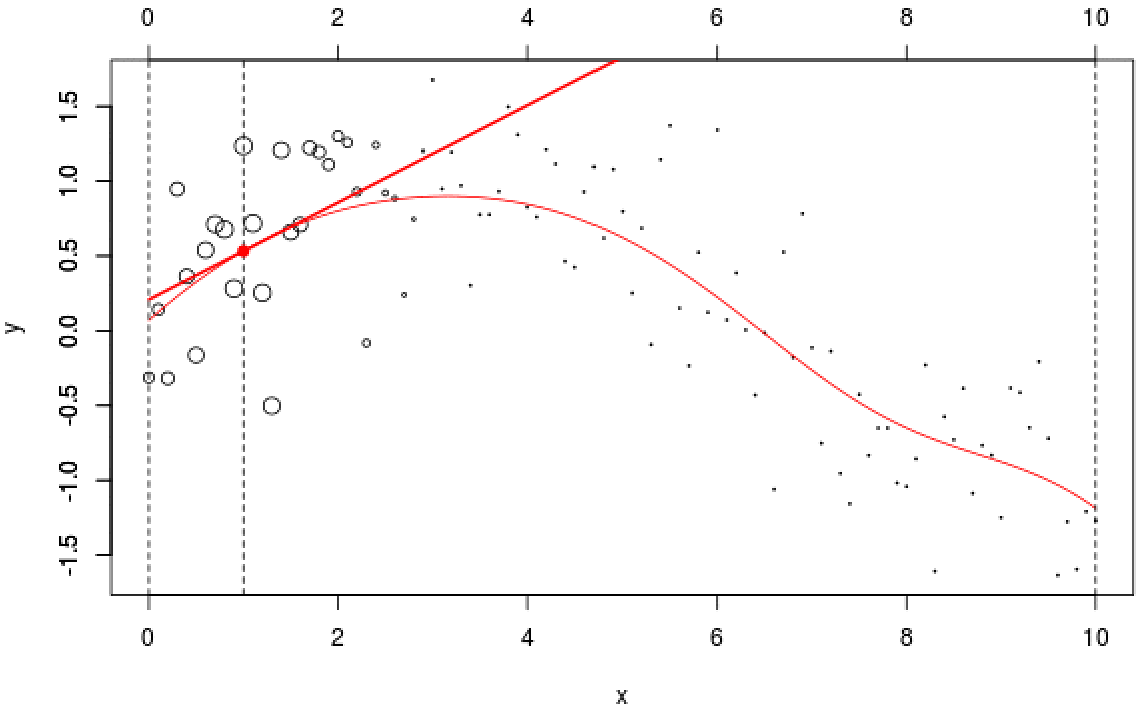


Nonlinear regression approaches required

Local linear regression

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

Local linear regression



Weighted polynomial regression.

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

Source: https://www.r-bloggers.com/some-heuristics-about-local-regression-and-kernel-smoothing/


Can we do more?

$$S_i = S_i^u + S_i^c$$

slide added by PowerPoint Labs

$$S_i = S_i^u + S_i^c$$

European Commission

Separating correlation from aggregation

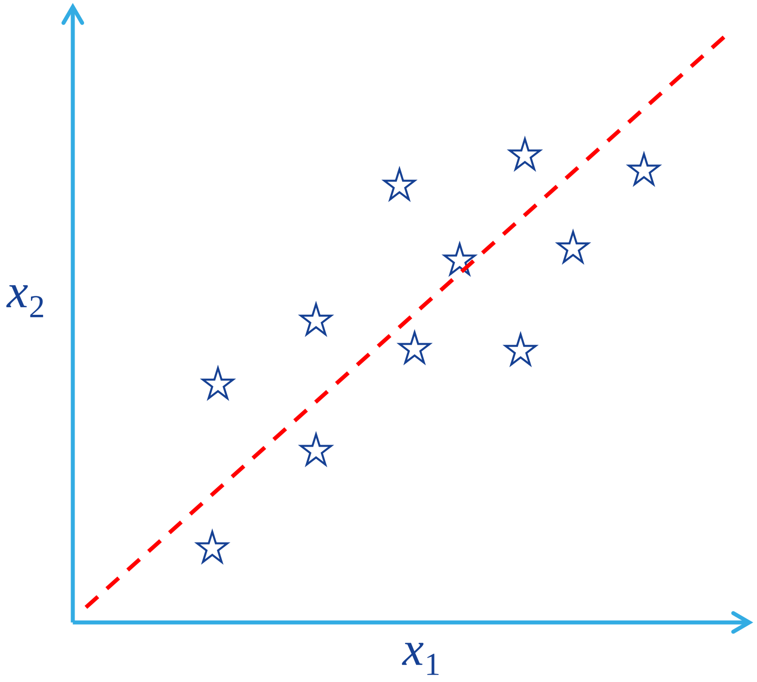
$$S_i = S_i^u + S_i^c$$

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

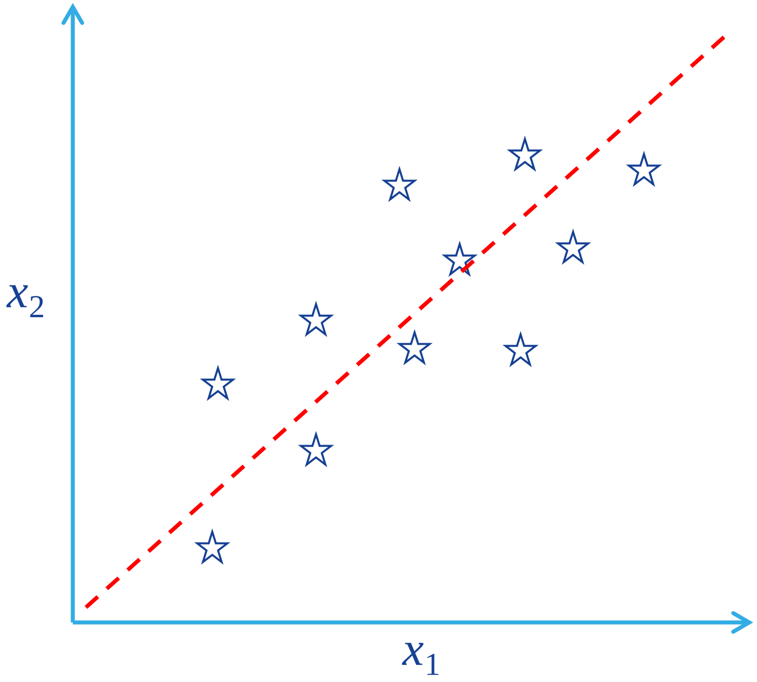
European Commission

Separating correlation from aggregation



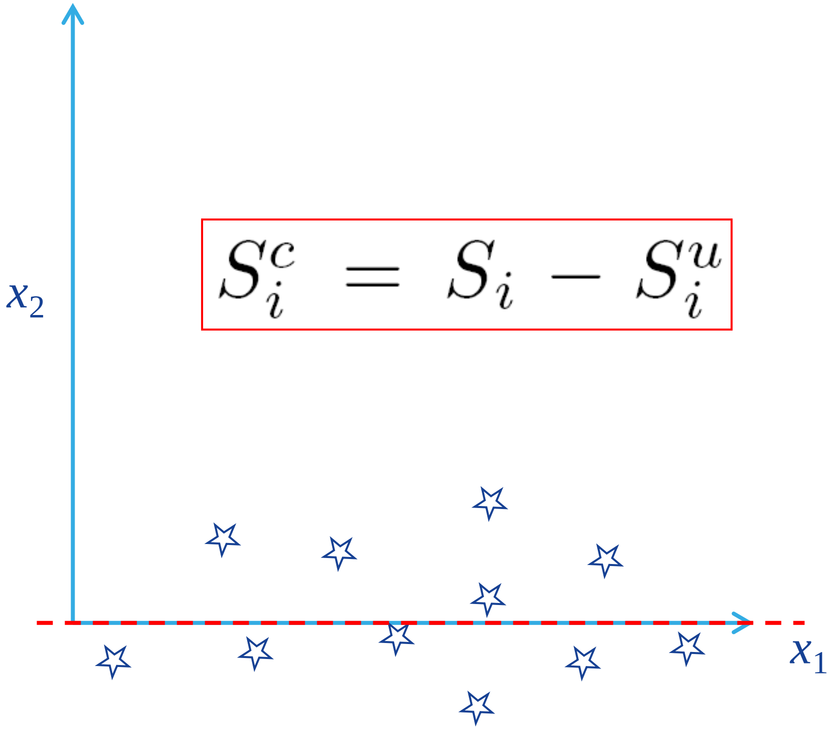
We regress x_i on to the remaining variables $x_{\sim i}$

Separating correlation from aggregation



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

Separating correlation from aggregation



$$\boxed{S_i^c = S_i - S_i^u}$$

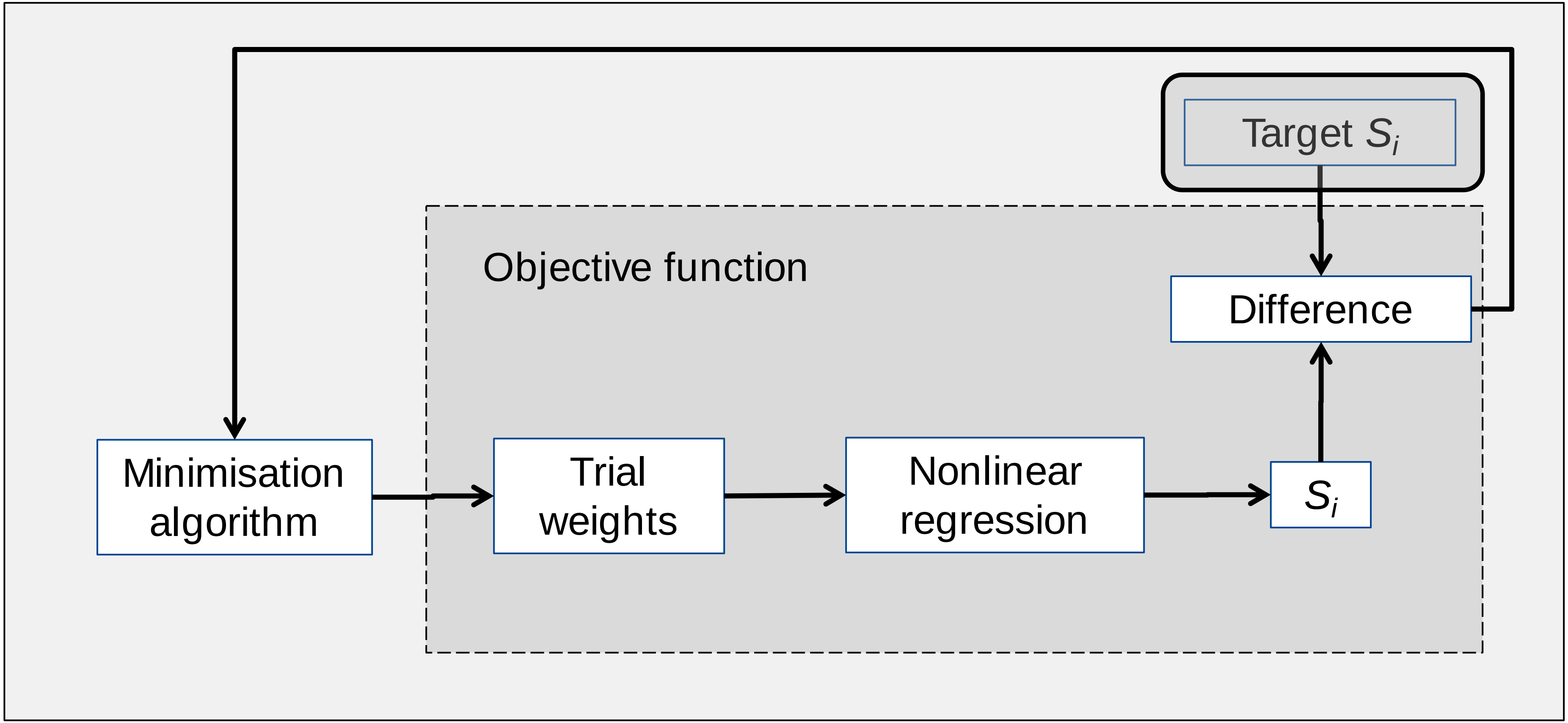
We regress x_i on to the remaining variables $x_{\sim i}$

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

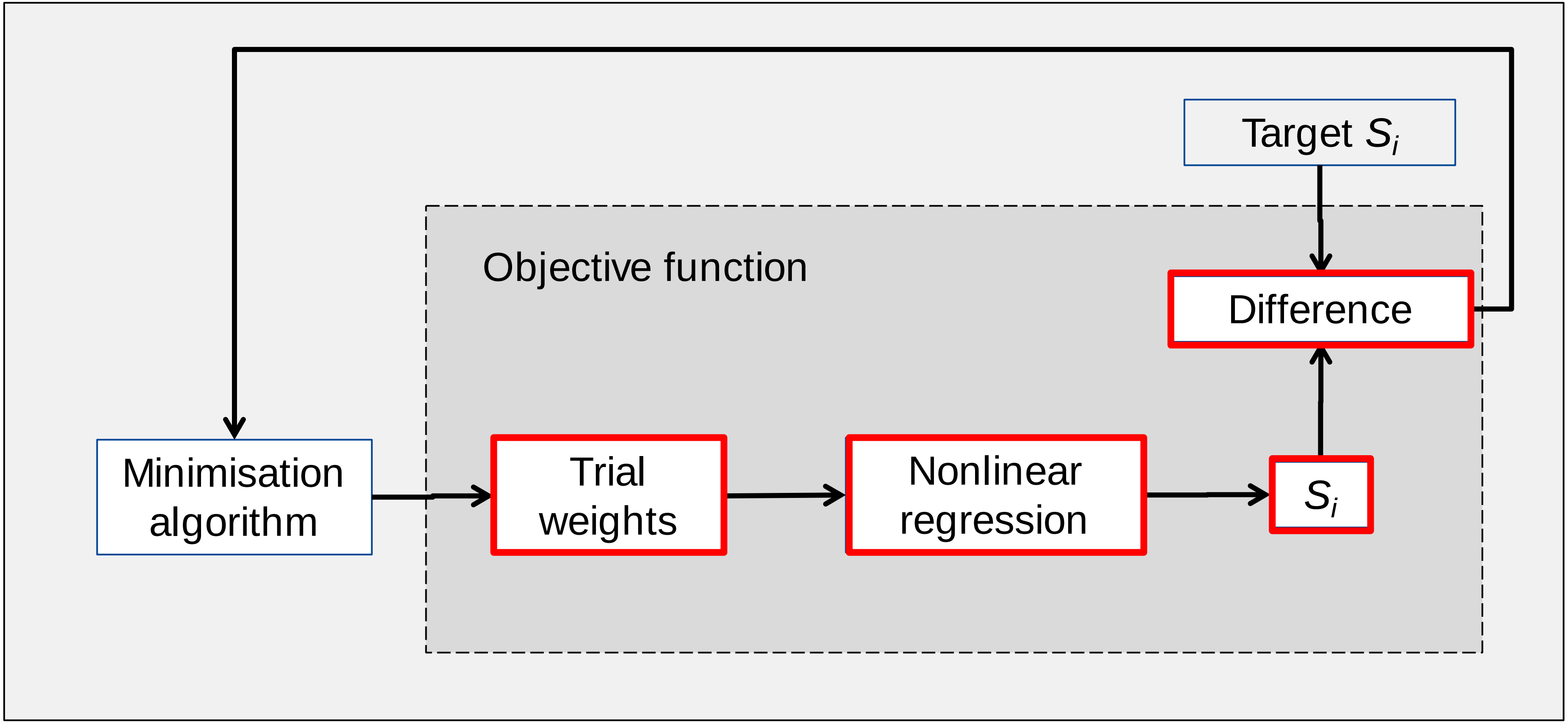
$$\boxed{\hat{z}_i = x_i - \hat{x}_i = x_i - \beta_0 + \sum_{l \neq i}^{d} \hat{\beta}_l x_l}$$

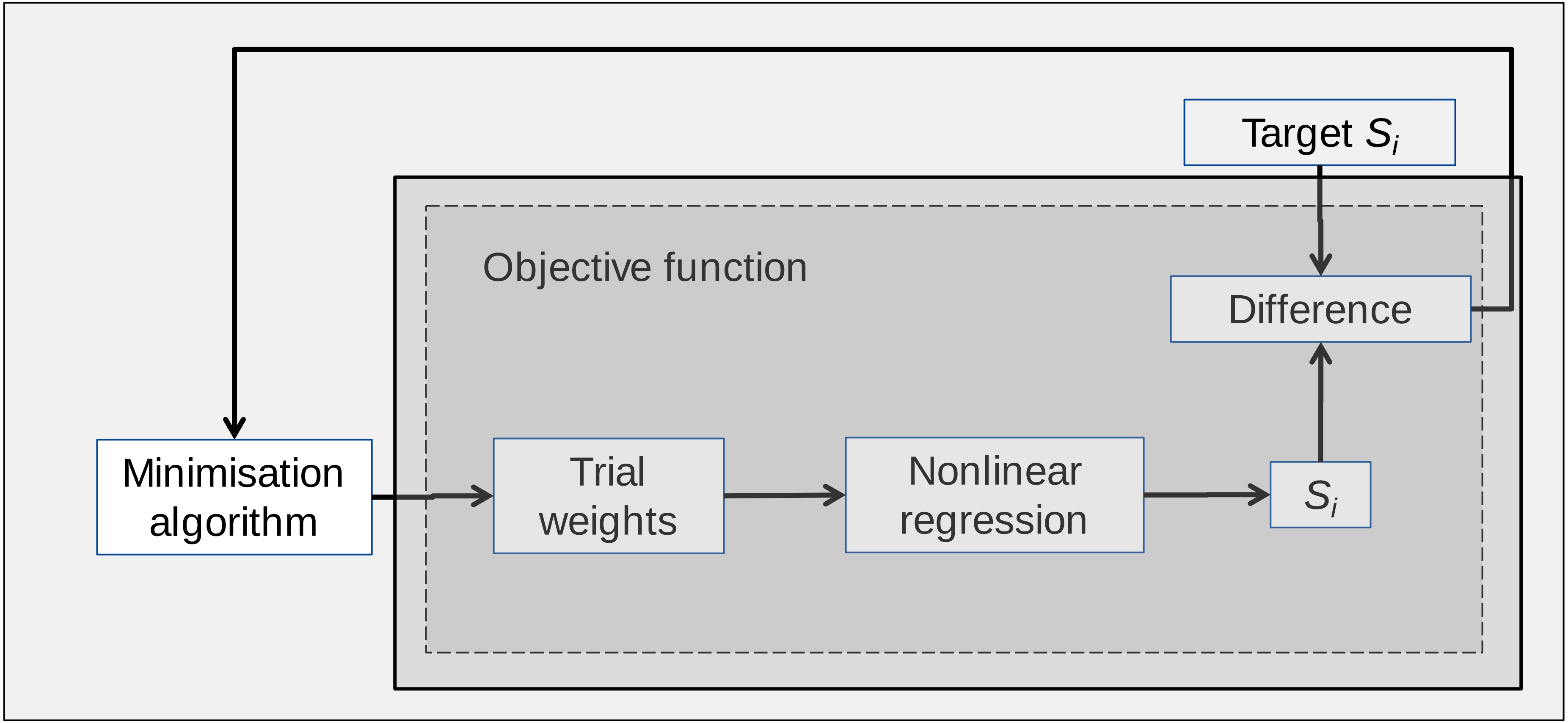
Now regress y on $\hat{z}_i$ (using NL regression) to get uncorrelated part.

Can we tune the weights?

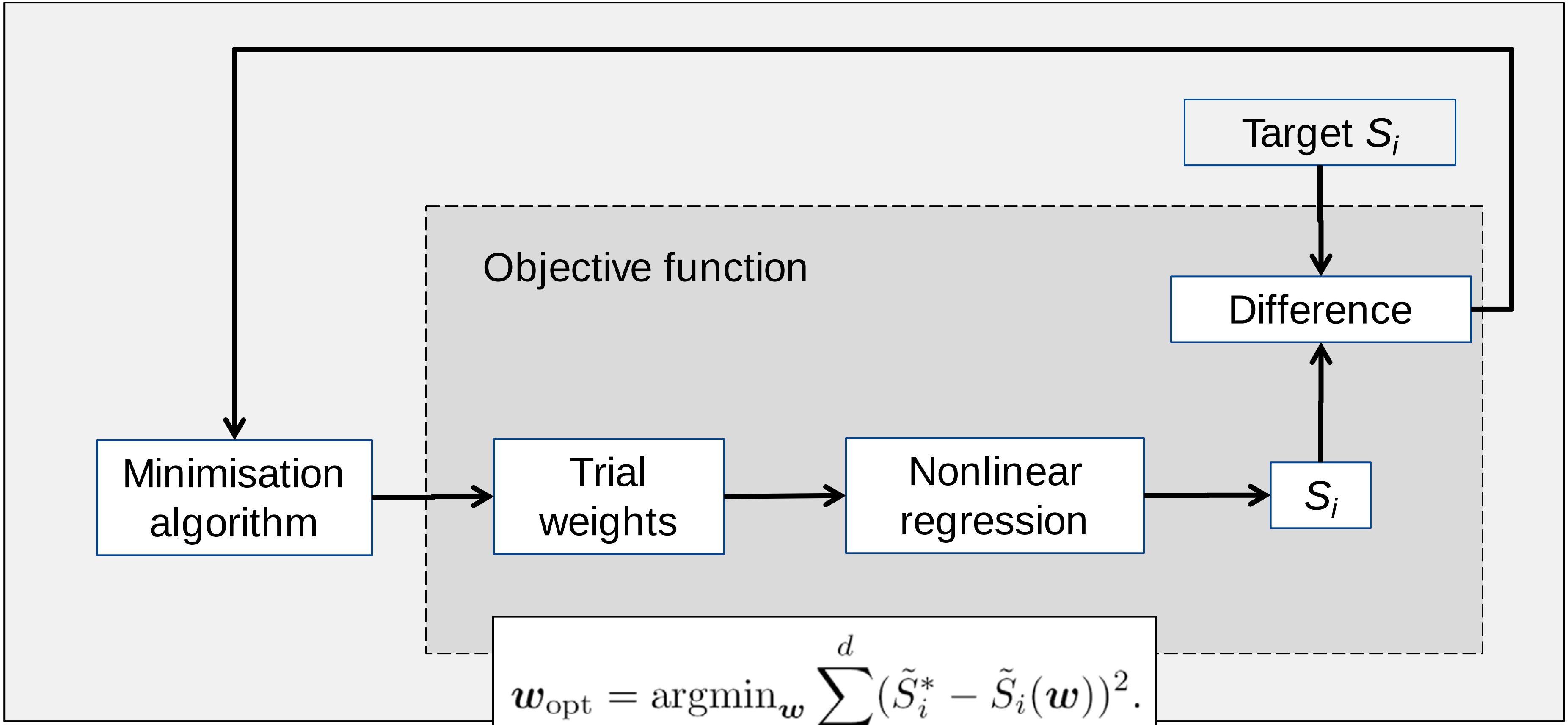


European Commission


Target S_i
Objective function
Difference
Minimisation algorithm
Trial weights
Nonlinear regression
S_i



European Commission



$$\boldsymbol{w}_{\text{opt}} = \text{argmin}_{\boldsymbol{w}} \sum_{i=1}^{d} (\tilde{S}_i^* - \tilde{S}_i(\boldsymbol{w}))^2.$$

Back to the beginning

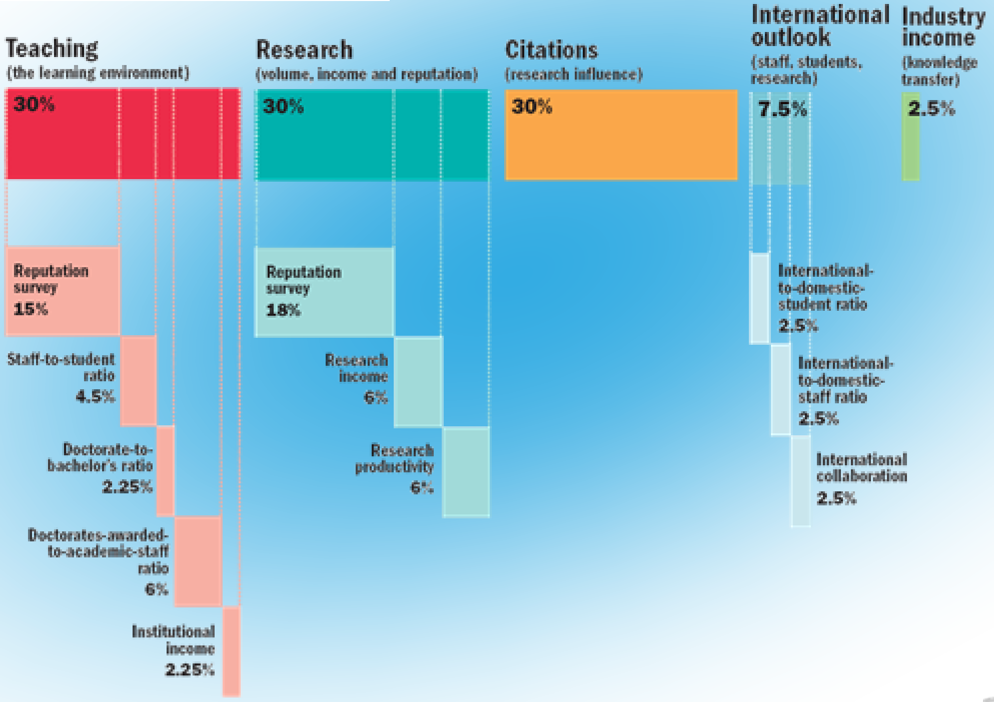
TIMES HIGHER EDUCATION WORLD UNIVERSITY RANKINGS

QS WORLD UNIVERSITY RANKINGS®

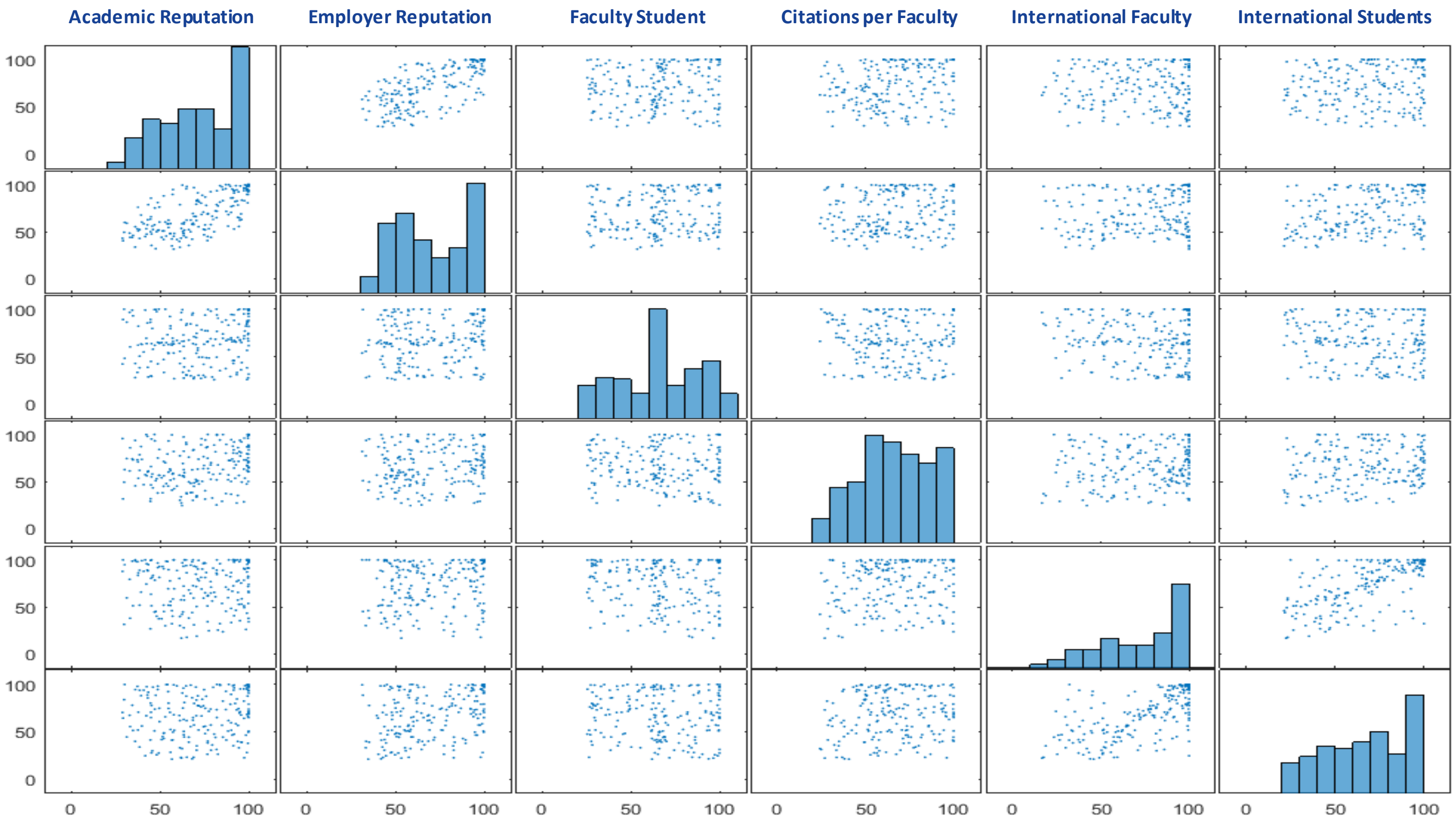
QS WORLD UNIVERSITY RANKINGS®

Academic Reputation	Employer Reputation	Faculty/ Student Ratio	Citations per Faculty	International Faculty	International Students
40%	10%	20%	20%	5%	5%
Opinion-based survey	Opinion-based survey	Teacher/ student ratio	Citations divided by faculty	Proportion of international faculty members	Proportion of international students

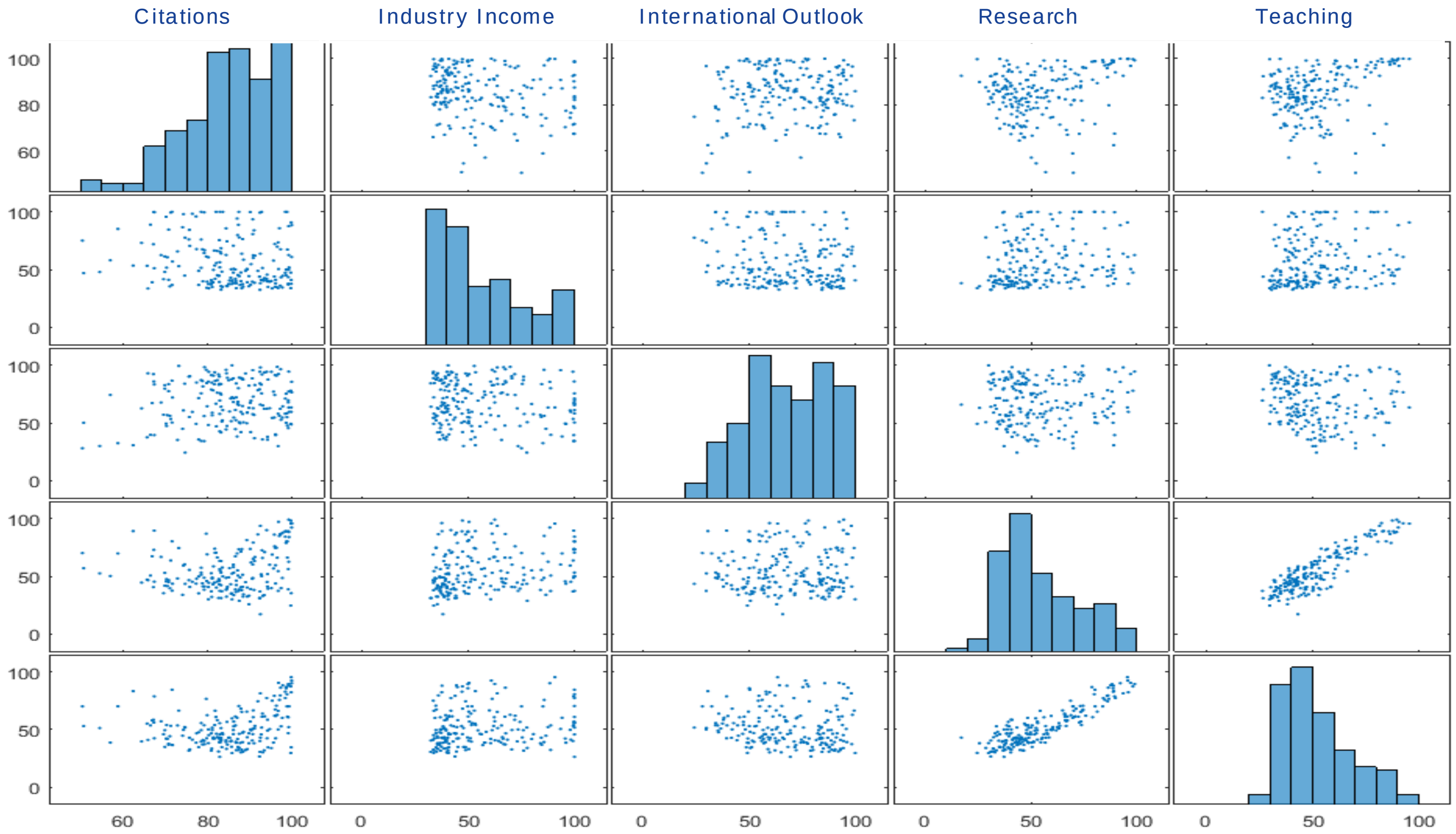
Academic Reputation	Employer Reputation	Faculty/ Student Ratio	Citations per Faculty	International Faculty	International Students
40%	10%	20%	20%	5%	5%
Opinion-based survey	Opinion-based survey	Teacher/ student ratio	Citations divided by faculty	Proportion of international faculty members	Proportion of international students

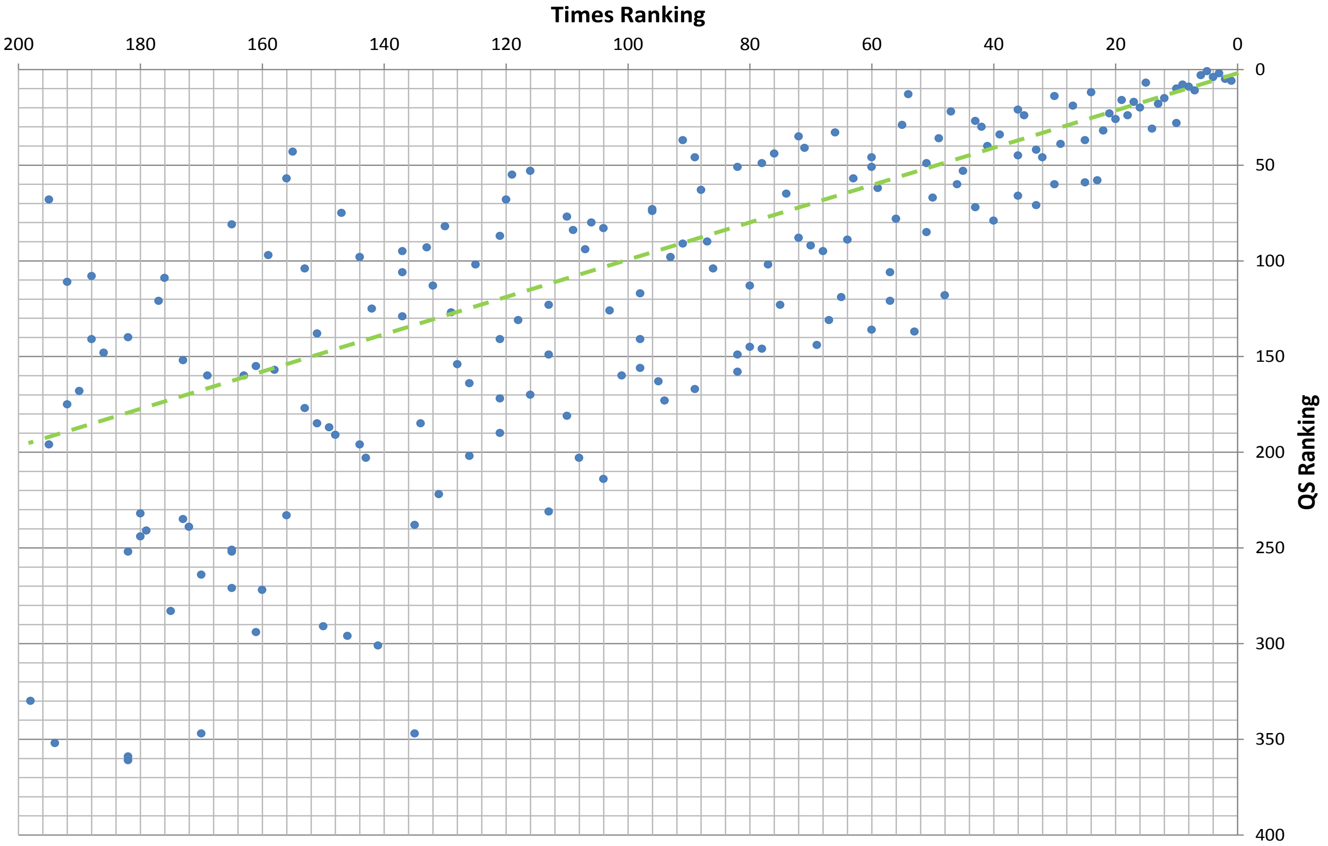

Teaching
(the learning environment)
30%
Reputation survey 15%
Staff-to-student ratio 4.5%
Doctorate-to-bachelor's ratio 2.25%
Doctorates-awarded-to-academic-staff ratio 6%
Institutional income 2.25%
Research
(volume, income and reputation)
30%
Reputation survey 18%
Research income 6%
Research productivity 6%
Citations
(research influence)
30%
International outlook
(staff, students, research)
7.5%
International-to-domestic-student ratio 2.5%
International-to-domestic-staff ratio 2.5%
International collaboration 2.5%
Industry income
(knowledge transfer)
2.5%

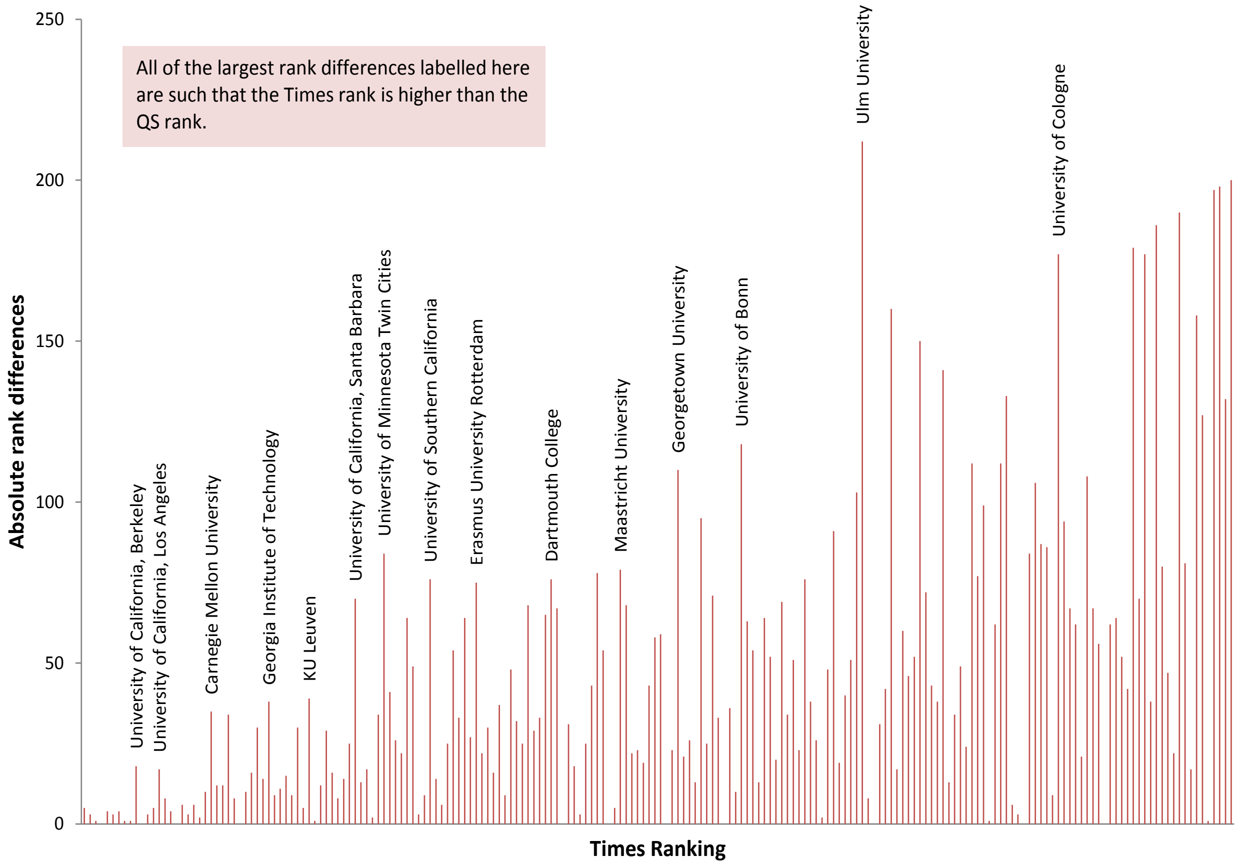
QS input data



Times input data

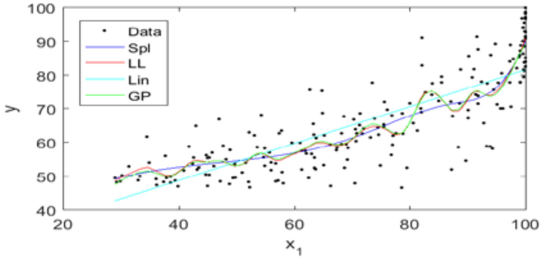


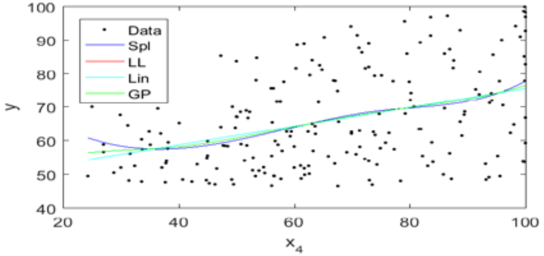

Times Ranking
200
180
160
140
120
100
80
60
40
20
0
QS Ranking
0
50
100
150
200
250
300
350
400

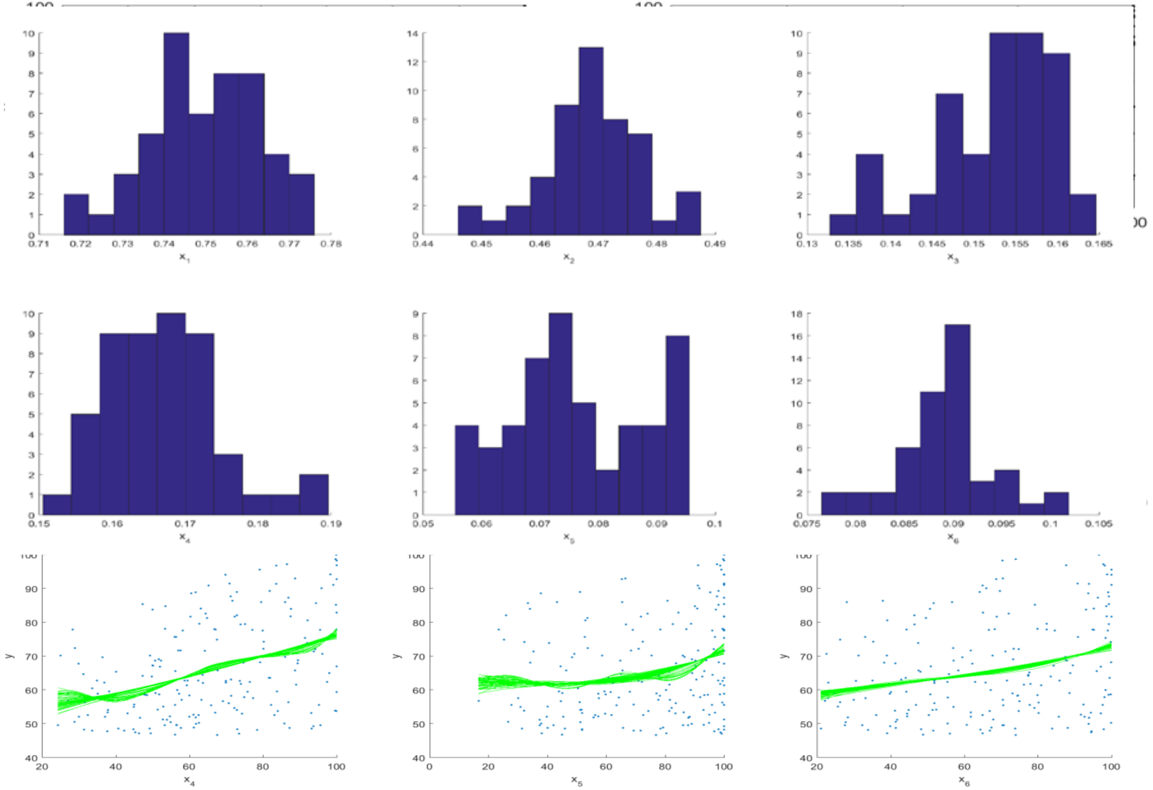


European Commission

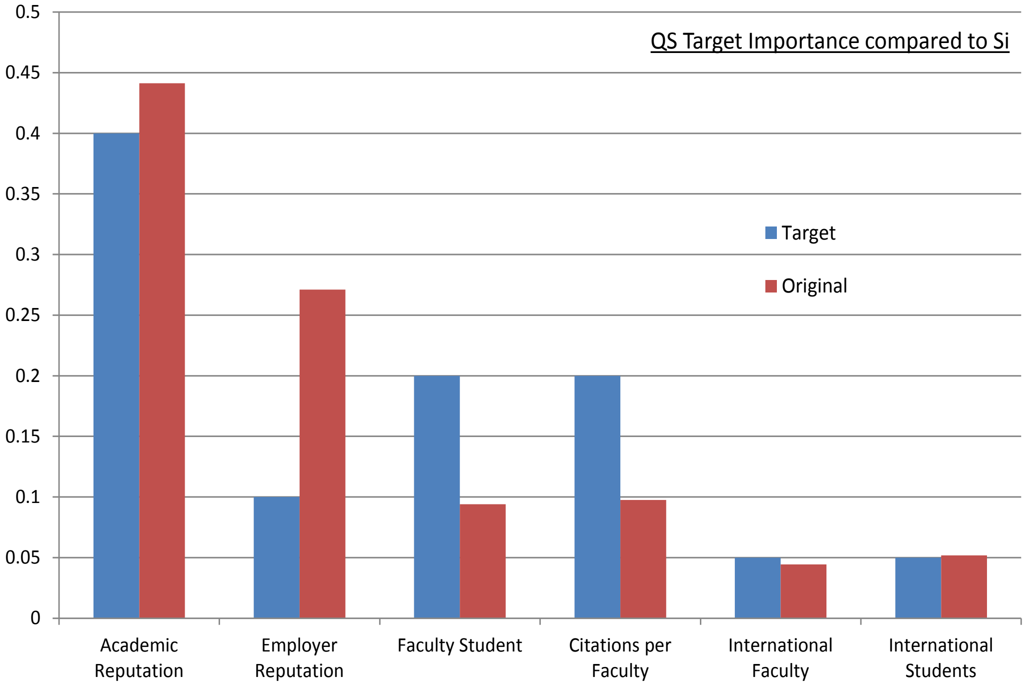
QS Regression



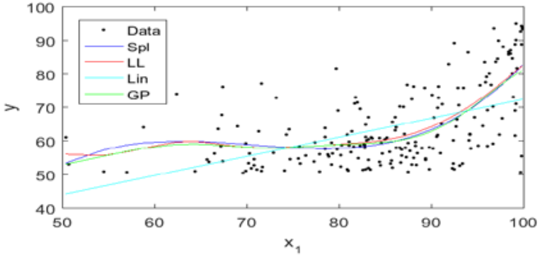


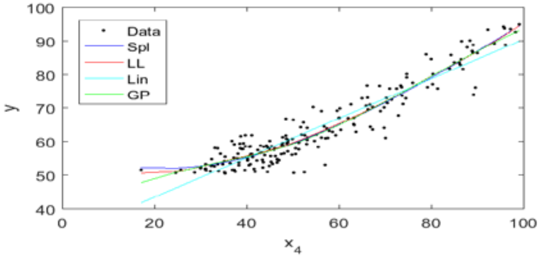


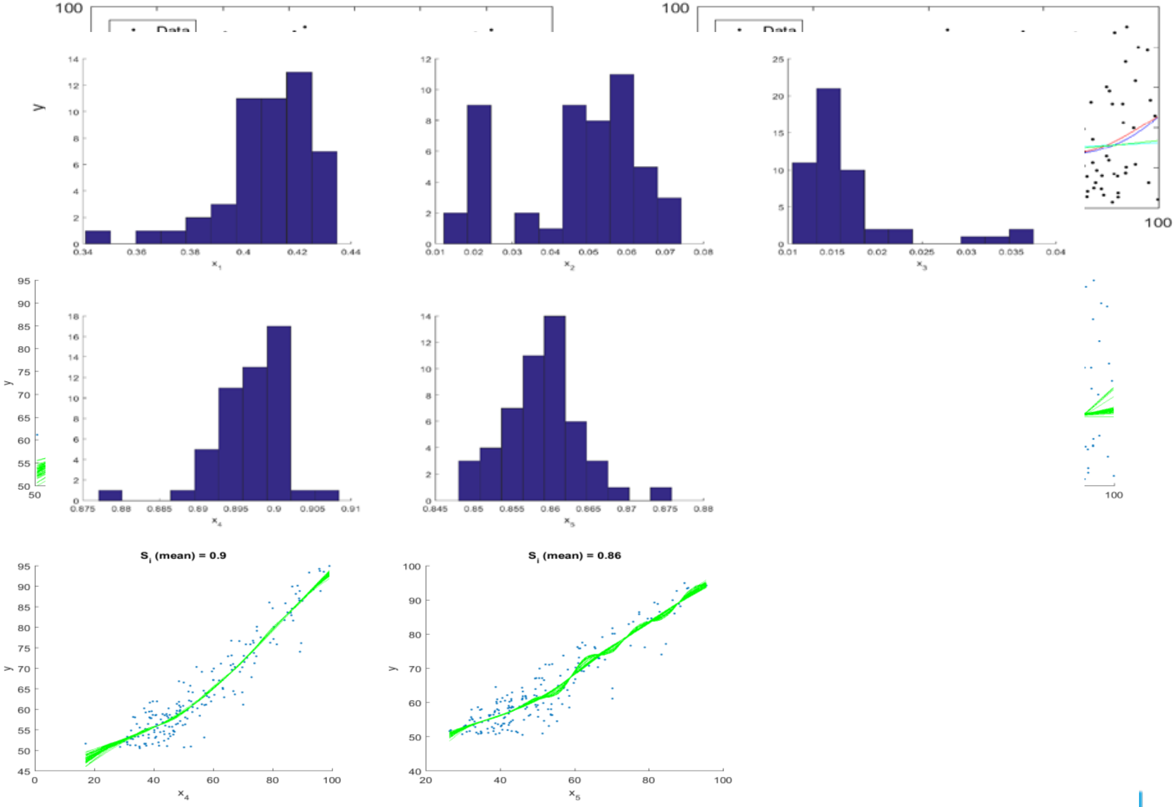
x1	Academic Reputation
x2	Employer Reputation
x3	Faculty Student
x4	Citations per Faculty
x5	International Faculty
x6	International Students


QS Target Importance compared to Si
Target
Original
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Academic Reputation
Employer Reputation
Faculty Student
Citations per Faculty
International Faculty
International Students

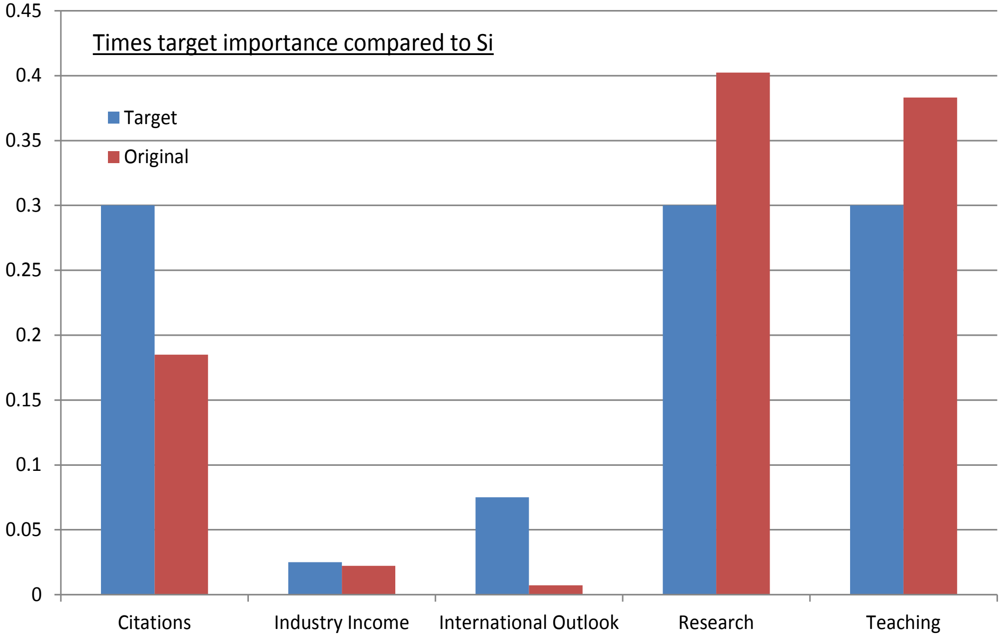
Times Regression



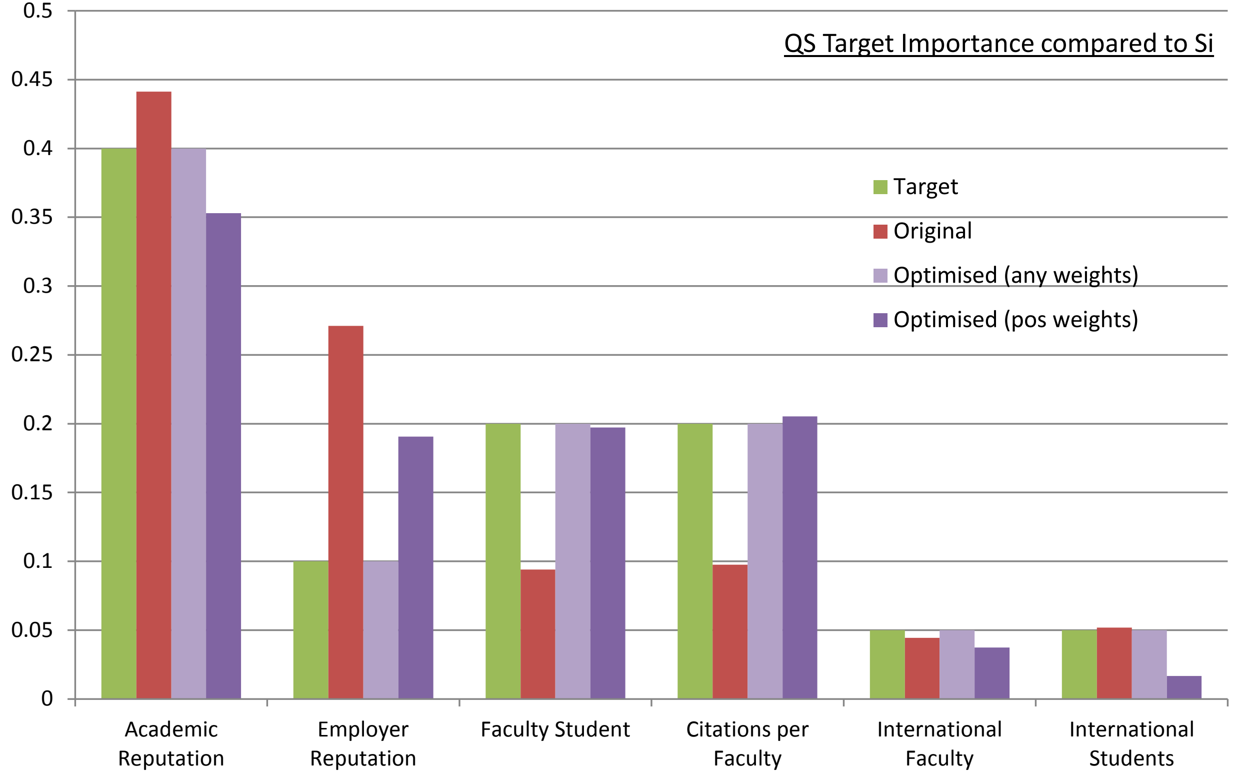




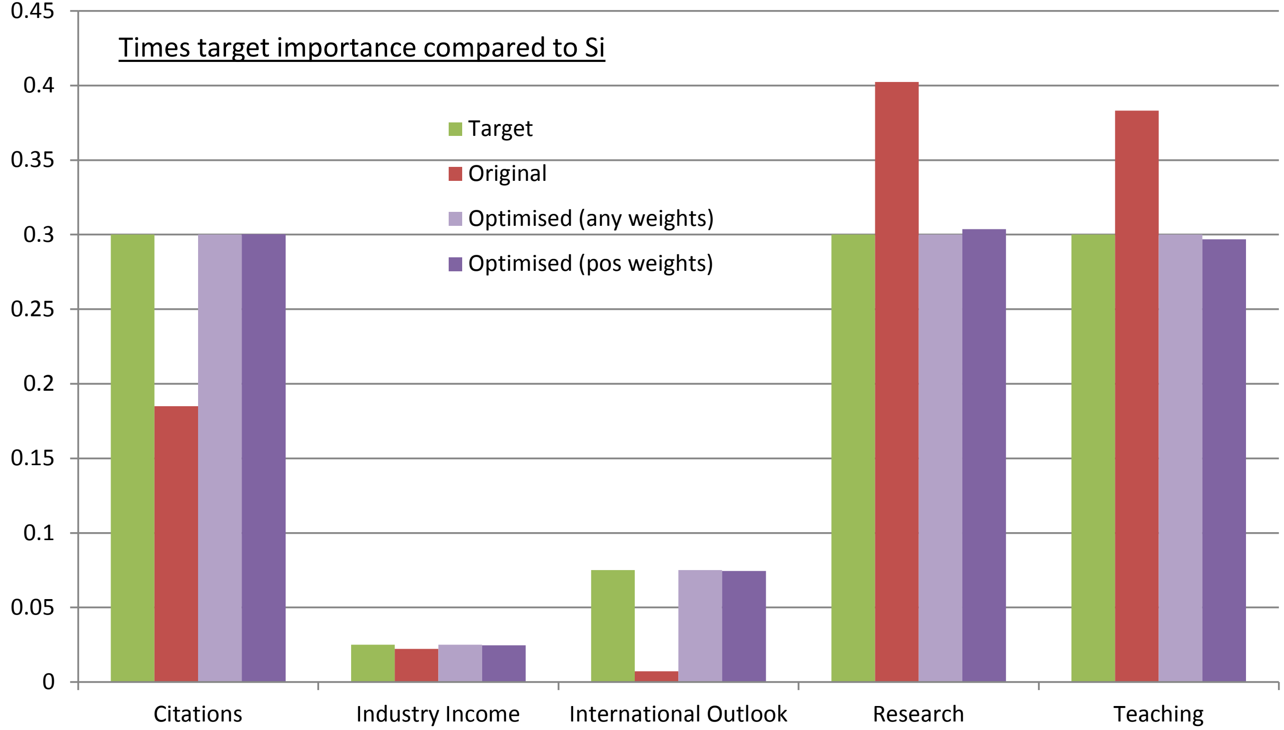
x1	Citations
x2	Industry Income
x3	International Outlook
x4	Research
x5	Teaching



European Commission



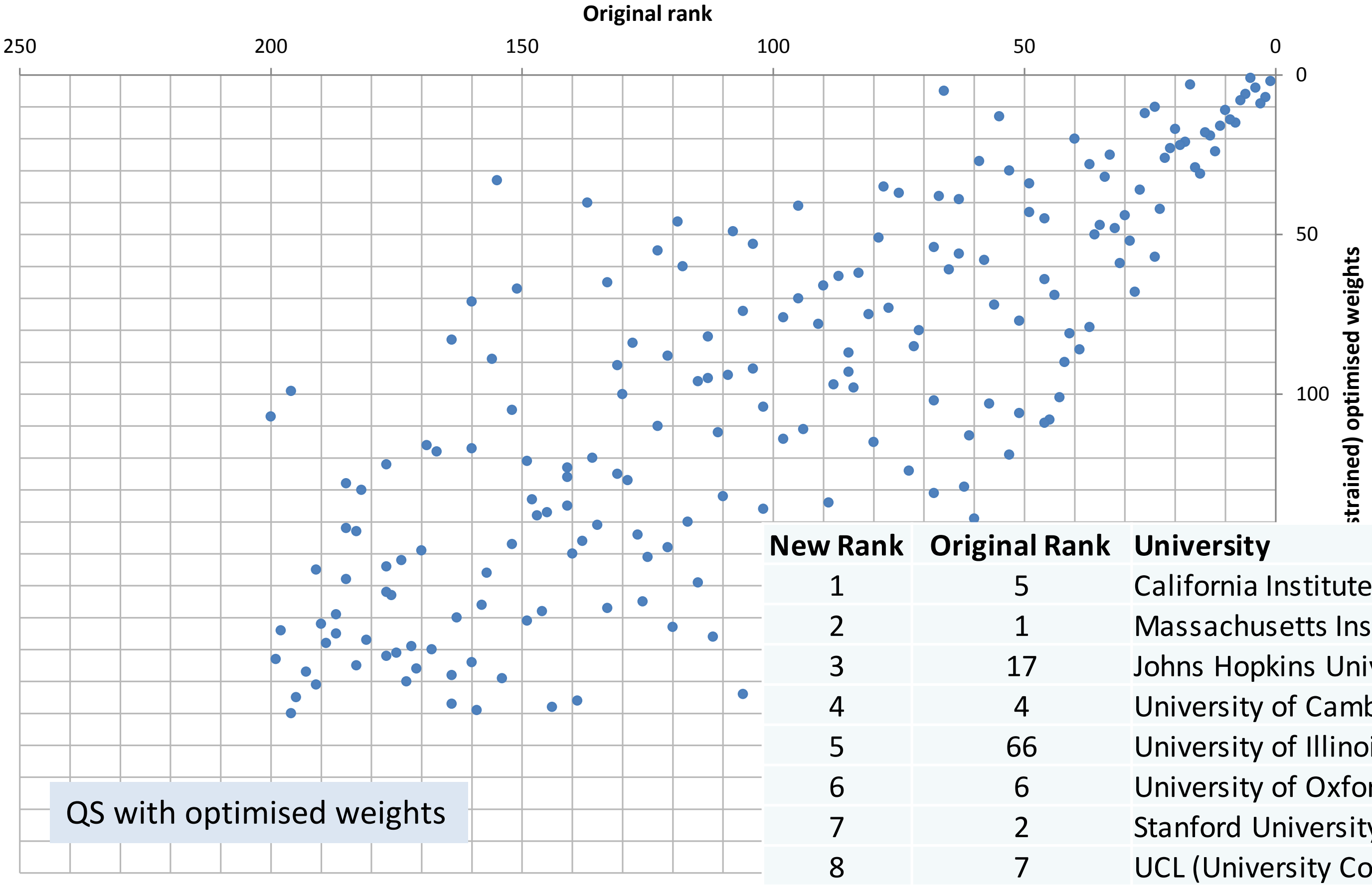
Original	0.4	0.1	0.2	0.2	0.05	0.05
Optimised	0.60	-0.37	0.32	0.28	-0.03	0.20
Opt (+)	0.32	0.00	0.30	0.31	0.07	0.00



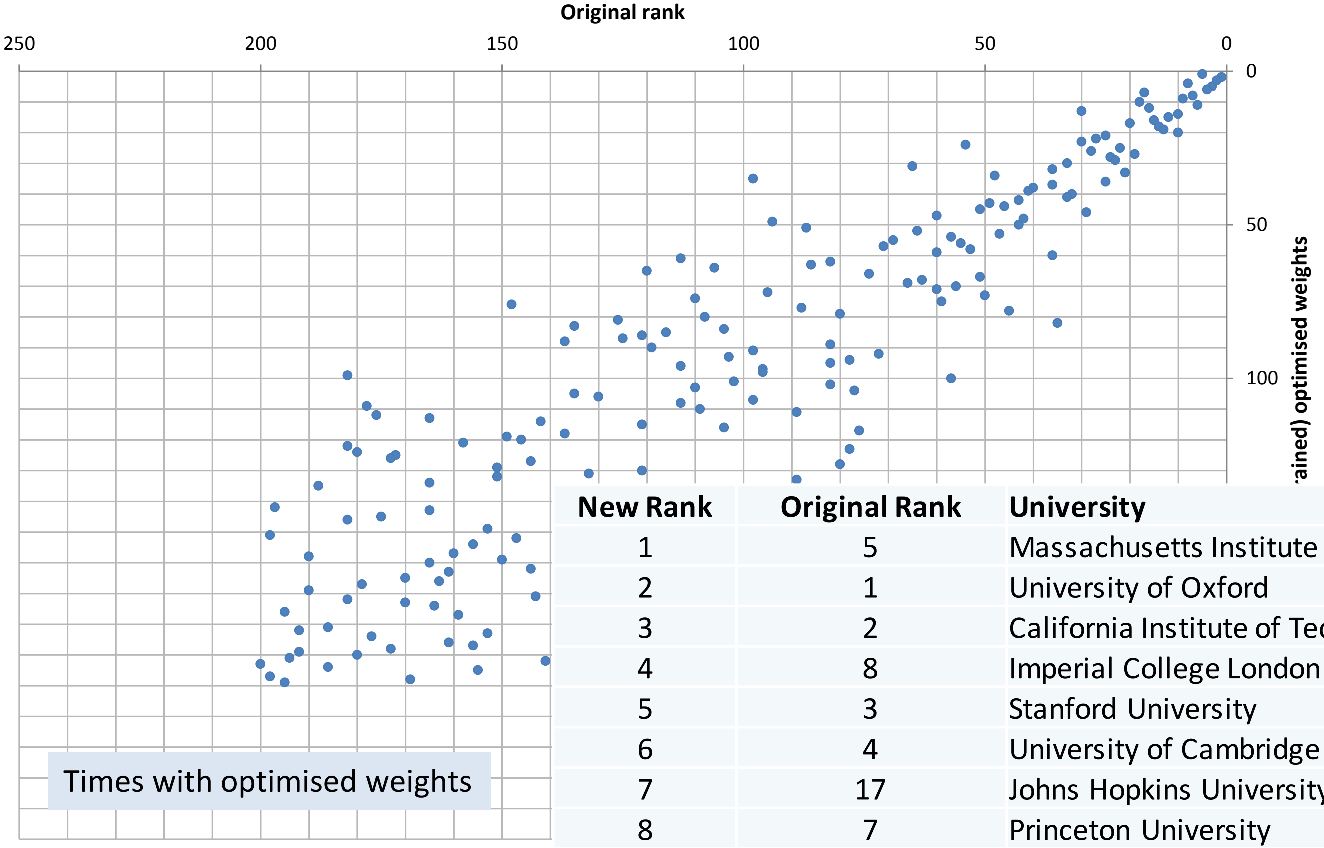
Original	0.30	0.03	0.08	0.30	0.30
Optimised	0.46	-0.08	0.19	0.03	0.39
Opt (+)	0.44	0.11	0.16	0.00	0.29

			S_i				S_i,u	S_i,c			
			Linear	Spline	Loc. Lin.	GP	Spline	Linear	Spline	Loc. Lin.	GP
QS	x1	Academic Reputation	0.67	0.75	0.78	0.75	0.27	0.40	0.48	0.51	0.48
	x2	Employer Reputation	0.45	0.48	0.48	0.46	0.07	0.38	0.41	0.40	0.39
	x3	Faculty Student	0.14	0.17	0.15	0.16	0.12	0.02	0.05	0.03	0.04
	x4	Citations per Faculty	0.16	0.17	0.16	0.17	0.08	0.08	0.09	0.08	0.08
	x5	International Faculty	0.06	0.10	0.09	0.08	0.04	0.01	0.05	0.05	0.03
	x6	International Students	0.09	0.10	0.08	0.09	0.07	0.01	0.03	0.01	0.02
Times	x1	Citations	0.27	0.43	0.43	0.41	0.10	0.17	0.34	0.33	0.32
	x2	Industry Income	0.02	0.05	0.04	0.05	0.01	0.01	0.05	0.04	0.04
	x3	International Outlook	0.01	0.03	0.03	0.02	0.02	0.00	0.01	0.02	0.00
	x4	Research	0.88	0.90	0.90	0.90	0.04	0.84	0.86	0.86	0.86
	x5	Teaching	0.85	0.86	0.86	0.85	0.05	0.80	0.81	0.82	0.81

- 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

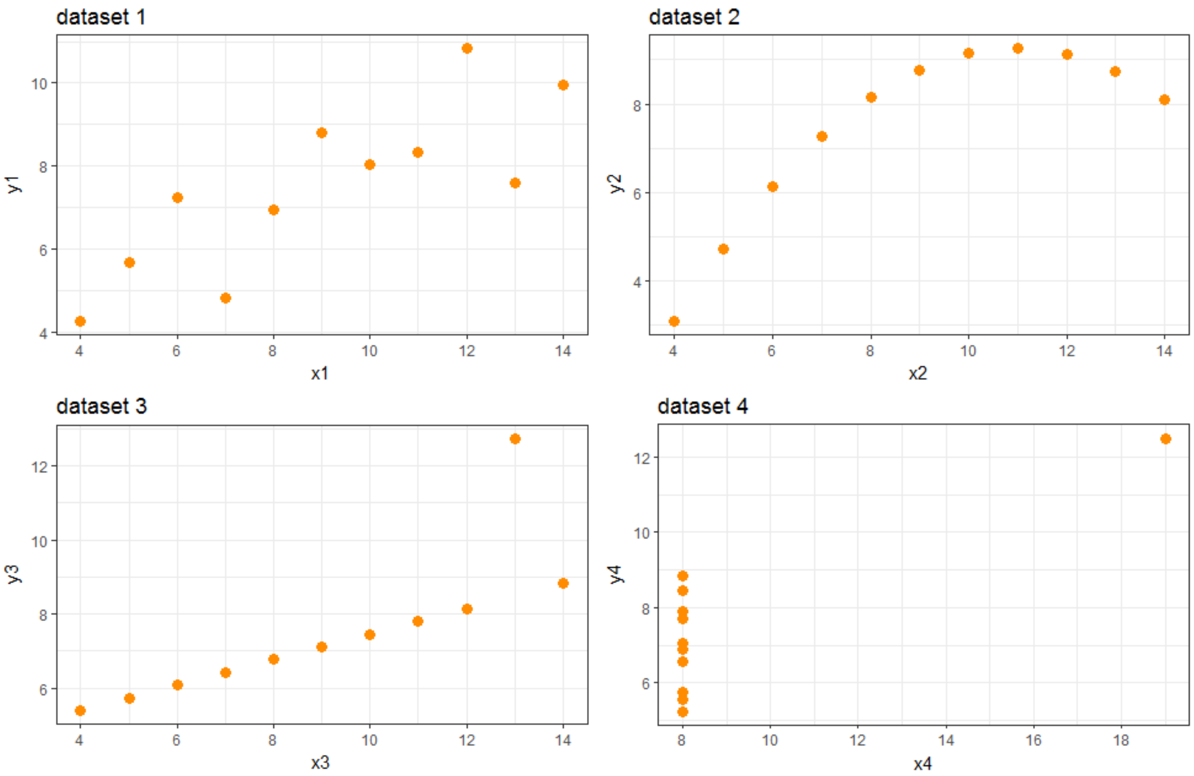


New Rank	Original Rank	University
1	5	California Institute of Technology (Caltech)
2	1	Massachusetts Institute of Technology (MIT)
3	17	Johns Hopkins University
4	4	University of Cambridge
5	66	University of Illinois at Urbana-Champaign
6	6	University of Oxford
7	2	Stanford University
8	7	UCL (University College London)
9	3	Harvard University
10	24	Duke University



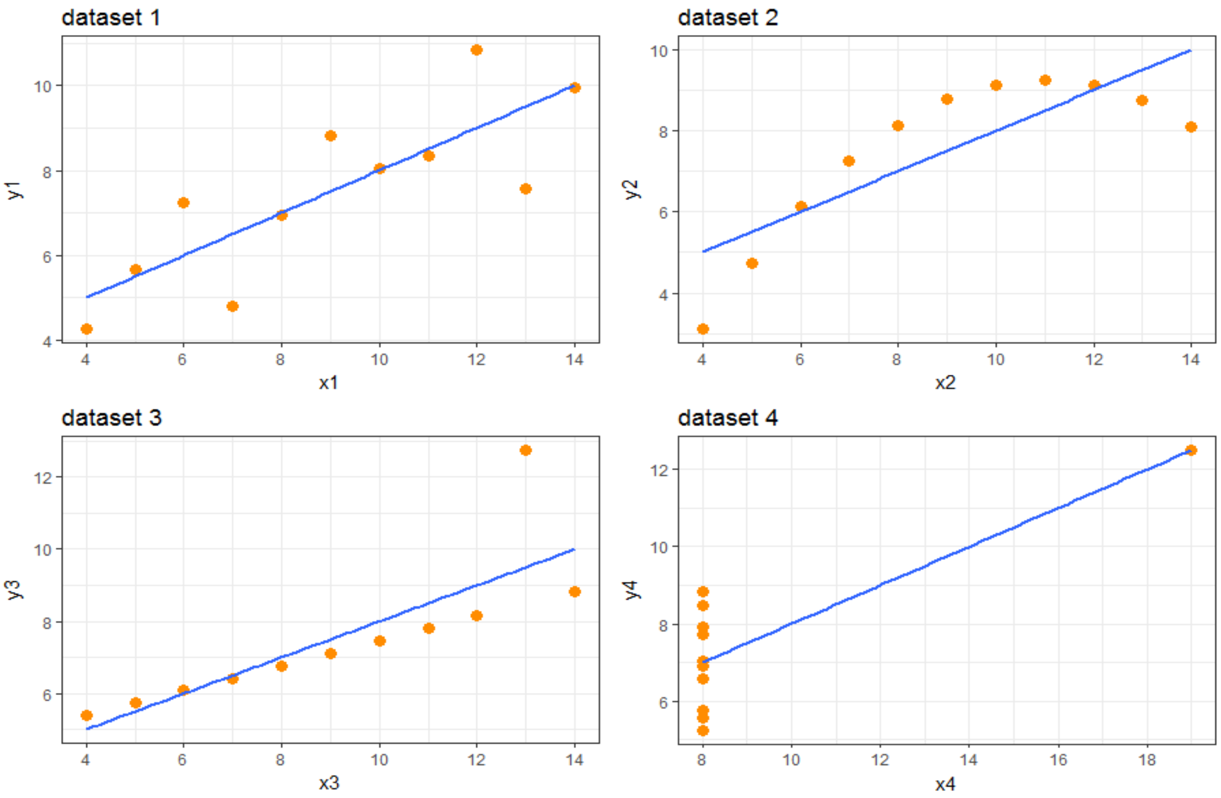
New Rank	Original Rank	University
1	5	Massachusetts Institute of Technology
2	1	University of Oxford
3	2	California Institute of Technology
4	8	Imperial College London
5	3	Stanford University
6	4	University of Cambridge
7	17	Johns Hopkins University
8	7	Princeton University
9	9	ETH Zurich
10	18	Duke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

A final thought

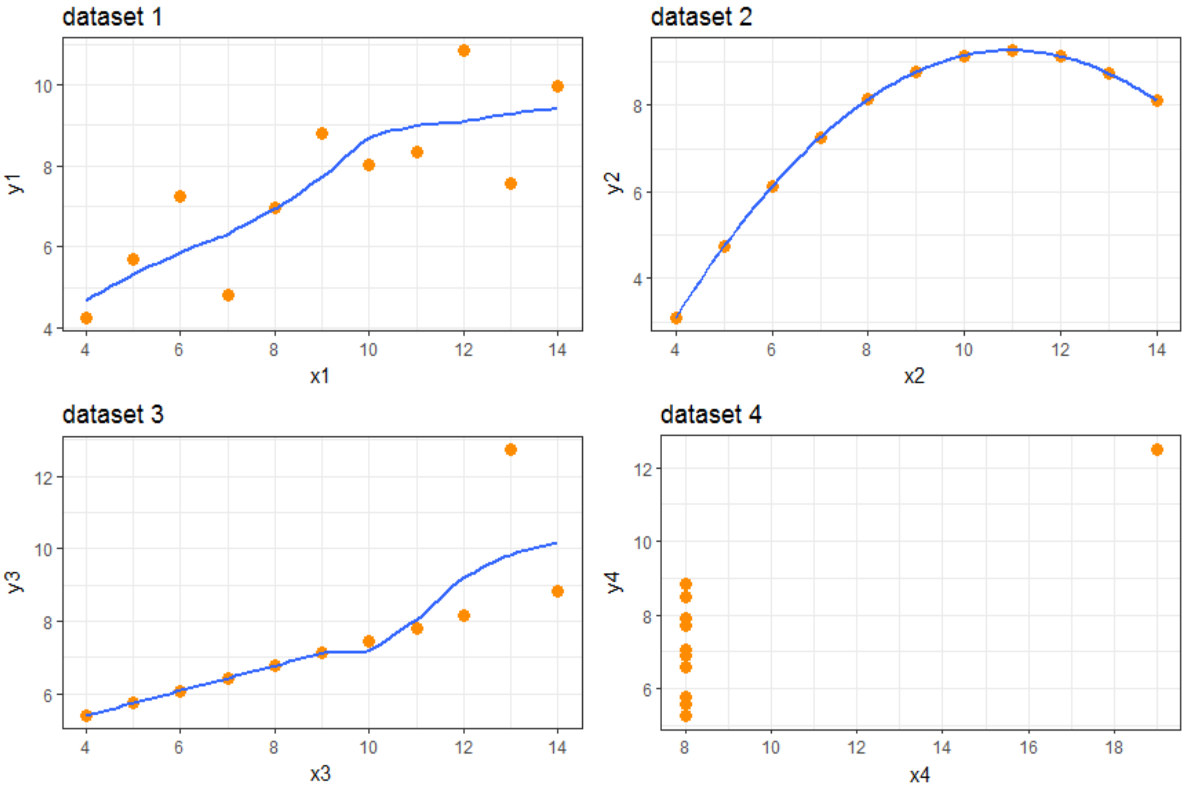


European Commission

A final thought

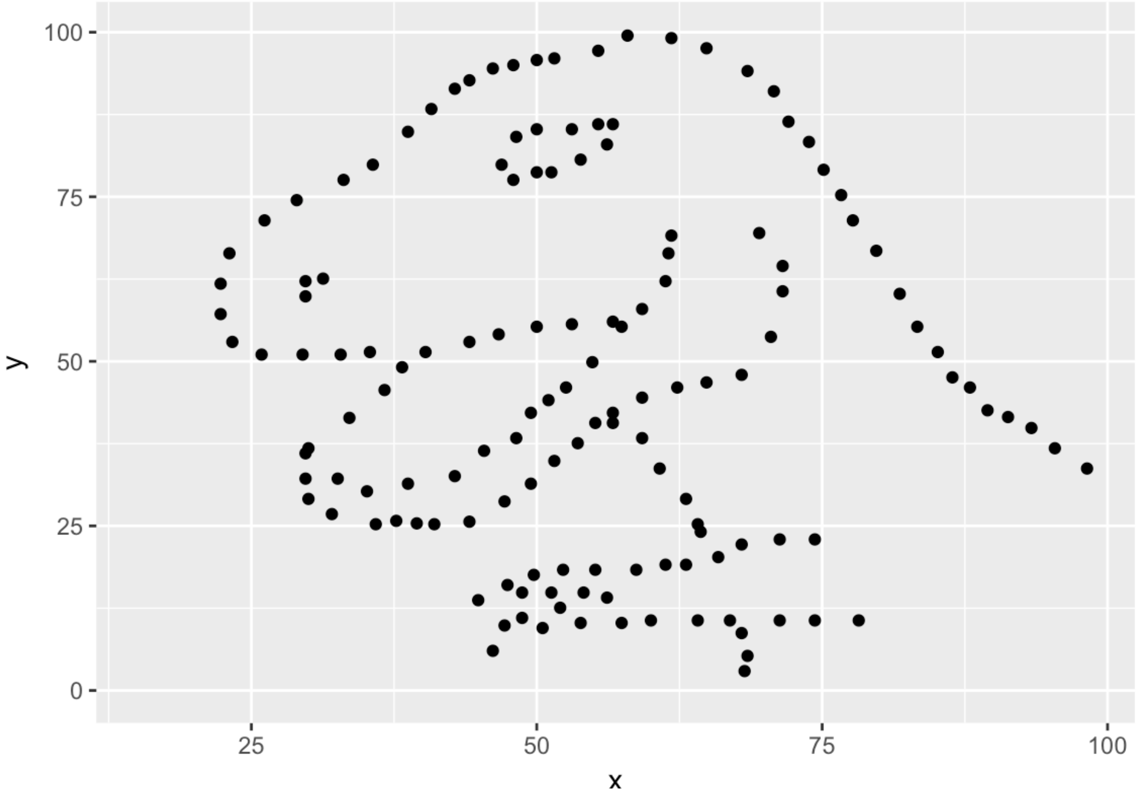


A final thought



European Commission

The Datasaurus Dozen



https://CRAN.R-project.org/package=datasauRus

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

All this can be found in recent paper:

Becker, W., Saisana, M., Paruolo, P., & Vandecasteele, I. (2017). Weights and importance in composite indicators: Closing the gap. *Ecological Indicators*, *80*, 12-22.

European Commission

Takeaway

Weights don’t equal importance in composite indicators.

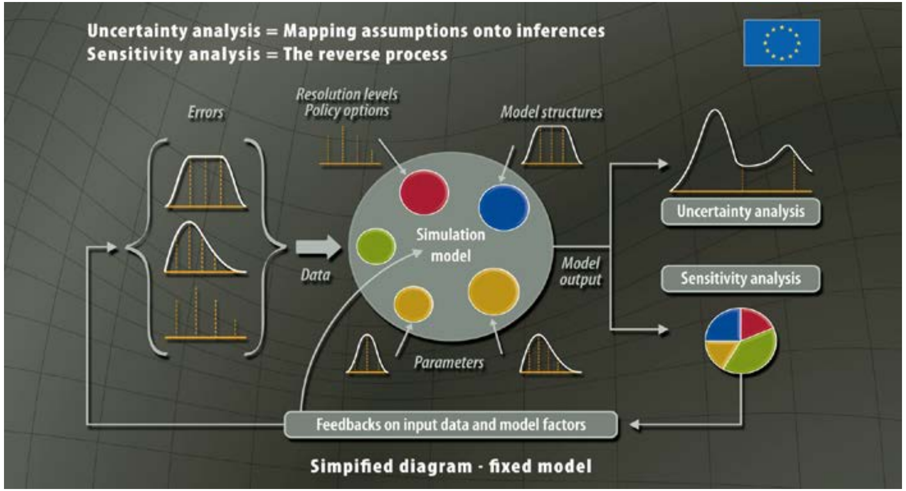
Open Questions

Does importance = S_i?

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

Part 2: Robustness and uncertainty analysis

Uncertainty and Sensitivity Analysis



Uncertainty analysis
How uncertain is the output (CI scores) given the uncertainty in the input (assumptions made in it's construction)?

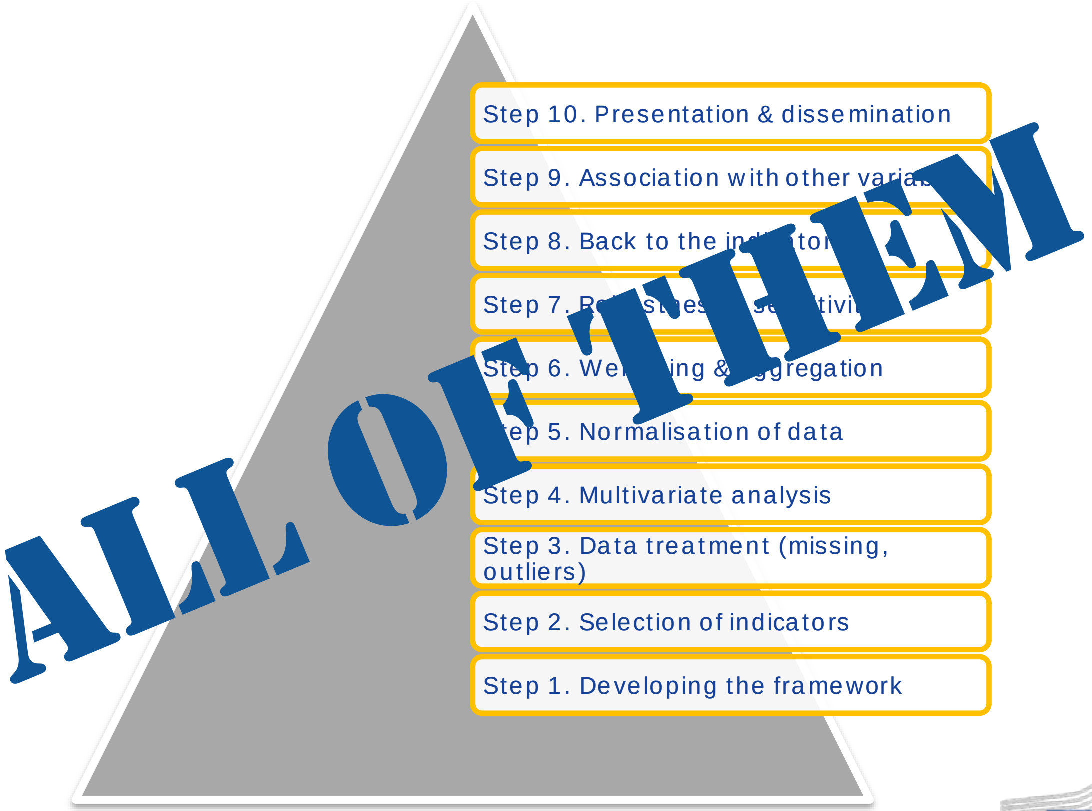
Sensitivity analysis
How much uncertainty is caused by each assumption?

We need:

- To quantify the uncertainty in our assumptions (assign probabilities to alternative assumptions)
- To propagate this through our composite indicator (Monte Carlo)
- To quantify/visualise uncertainty in the scores of our composite indicator (confidence intervals, pdfs, scatter plots, sensitivity analysis tools)

European Commission

Which steps in the construction of a composite indicator are uncertain?



European Commission

Which steps in the construction of a composite indicator are uncertain?

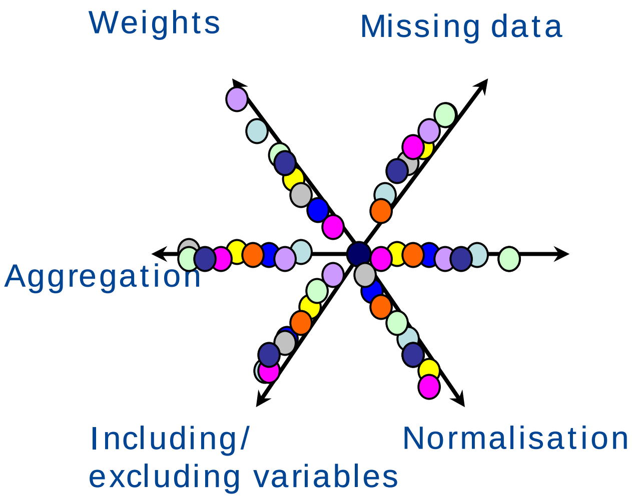
- Step 10. Presentation & dissemination
- Step 9. Association with other varia[illegible]
- Step 8. Back to the in[illegible]tor[illegible]
- Step 7. R[illegible]st[illegible]es[illegible]s[illegible]tivi[illegible]
- Step 6. We[illegible]ing & [illegible]ggregation
- [illegible]ep 5. Normalisation of data
- Step 4. Multivariate analysis
- Step 3. Data treatment (missing, outliers)
- Step 2. Selection of indicators
- Step 1. Developing the framework

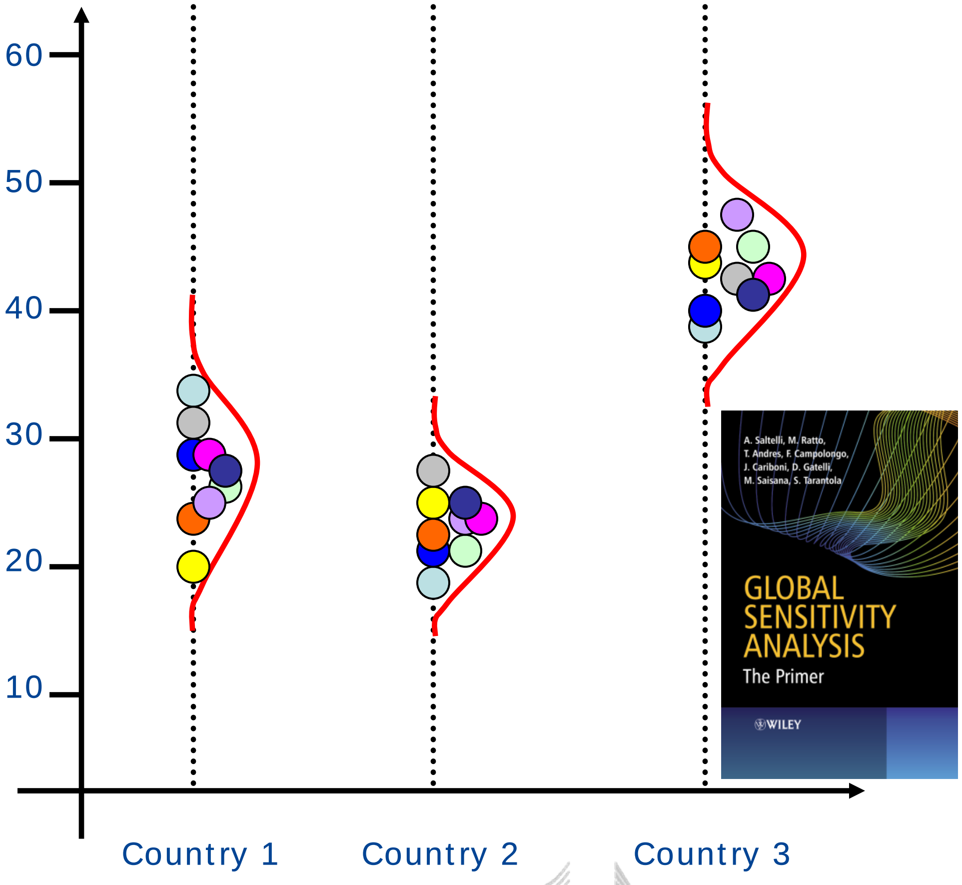
ALL OF THEM

Which steps in the construction of a composite indicator are uncertain?

Step 10. Presentation & dissemination	
Step 9. Association with other variables	
Step 8. Back to the indicators	
Step 7. Robustness & sensitivity	Even here: uncertain pdfs, limited exploration of assumptions...
Step 6. Weighting & aggregation	What weights? How to aggregate?
Step 5. Normalisation of data	How? Min/max, standarise, rank...
Step 4. Multivariate analysis	Which methods? Assume normality?
Step 3. Data treatment (missing, outliers)	How to impute? What is an outlier?
Step 2. Selection of indicators	The right ones? Irrelevant ones? Missing indicators?
Step 1. Developing the framework	What is the definition of "excellence"/"innovation"/etc...?

Space of alternatives





Model averaging: whenever a choice in the composite setting-up may not be strongly supported or if you may not trust one single model, we'll recommend you to use more models

it is essential to vary assumptions simultaneously to fully explore the “assumption space”.

C_1

Only child 1:

chaos = $f_1(C_1)$

C_2

Only child 2:

chaos = $f_2(C_2)$

Child 1 *plus* child 2:

Fighting, crying interactions, etc.

chaos = $f_1(C_1) + f_2(C_2) + f_{12}(C_1,C_2)$ $\boxed{\neq f_1(C_1) + f_2(C_2)}$

Testing the behaviour of one child at a time would not at all explore the full space of chaos.

Uncertainty analysis of composite indicators requires some programming—we use Matlab, but potentially moving to R.

Testable assumptions include:

- Normalisation method (A_1)
- Weighting method (A_2)
- Perturbations of weights (A_3)
- Set of indicators included (A_4)
- Data imputation method (…)
- Structure of composite (…)
- ++ anything you can program! (…)

$$\boxed{\text{scores} = CI(A_1, A_2, A_3, A_4, \ldots)}$$

Sample this function many times at random A_i values, and record the output each time

What do uncertainty and sensitivity analyses tell you?

→ *NOT to verify whether the two global university rankings are legitimate models to measure university performance*

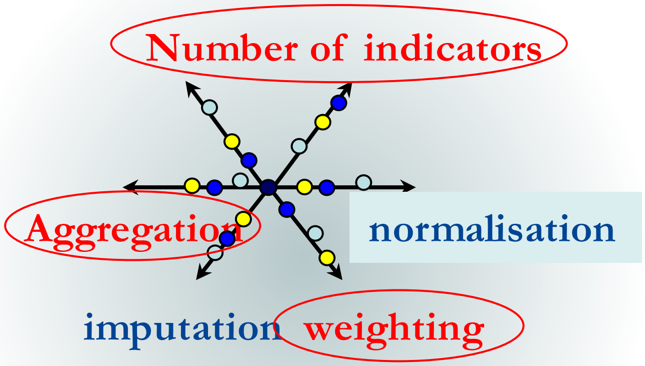
→ *To test whether the rankings and/or their associated inferences are robust or volatile with respect to changes in selected methodological assumptions within a plausible and legitimate range.*

Source: Saisana, D'Hombres, Saltelli, 2011, Research Policy 40, 165–177

European Commission

Back to university rankings

Activate simultaneously different sources of uncertainty that cover a wide spectrum of methodological assumptions



70 scenarios

Assumption	Alternatives
Number of indicators	▪ all six indicators included or one-at-time excluded (6 options)
Weighting method	▪ original set of weights, ▪ factor analysis, ▪ equal weighting, ▪ data envelopment analysis
Aggregation rule	▪ additive, ▪ multiplicative, ▪ Borda multi-criterion

Estimate the FREQUENCY of the university ranks obtained in the different simulations

European Commission

ARWU/SJTU

Legend:
Frequency lower 15%
Frequency between 15 and 30%
Frequency between 30 and 50%
Frequency greater than 50%
Note: Frequencies lower than 4% are not shown

Simulated rank range - SJTU 2008

	1-5	6-10	11-15	16-20	21-25	26-30	31-35	36-40	41-45	46-50	51-55	56-60	61-65	66-70	71-75	76-80	81-85	86-90	91-95	96-100	Original rank	
Harvard Univ	100																				1	USA
Stanford Univ	89	11																			2	USA
Univ California - Berkeley	97																				3	USA
Univ Cambridge	90	10																			4	UK
Massachusetts Inst Tech (MIT)	74	26																			5	USA
California Inst Tech	27	53	19																		6	USA
Columbia Univ	23	77																			7	USA
Princeton Univ		71	9	11	7																8	USA
Univ Chicago		51	34	13																	9	USA
Univ Oxford		99																			10	UK
Yale Univ		47	53																		11	USA
Cornell Univ		27	73																		12	USA
Univ California - Los Angeles		9	84	7																	13	USA
Univ California - San Diego			41	46	9																14	USA
Univ Pennsylvania		6	71	23																	15	USA
Univ Washington - Seattle			7	71	21																16	USA
Univ Wisconsin - Madison			27	70																	17	USA
Univ California - San Francisco			14	9	14		11		7	10						6				6	18	USA
Tokyo Univ			16	16	49	20															19	Japan
Johns Hopkins Univ			7	54	21	17															20	USA

- Harvard, Stanford, Berkley, Cambridge, MIT: top 5 in more than 75% of JRC simulations.
- Univ California: original rank 18th but could be ranked anywhere between the 6th and 100th position
- Impact of assumptions: much stronger for the middle ranked universities

European Commission

THE Ranking

TIMES HIGHER EDUCATION WORLD UNIVERSITY RANKINGS

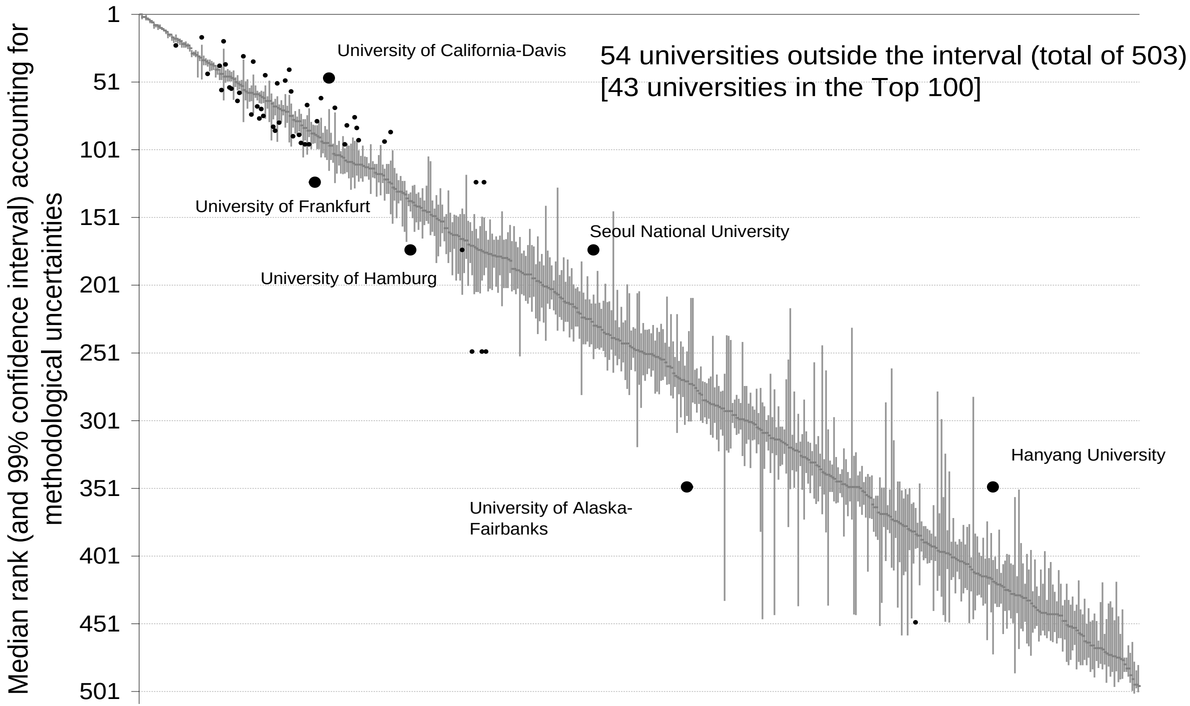
Legend:
Frequency lower 15%
Frequency between 15 and 30%
Frequency between 30 and 50%
Frequency greater than 50%
Note: Frequencies lower than 4% are not shown

Simulated rank range - THES 2008

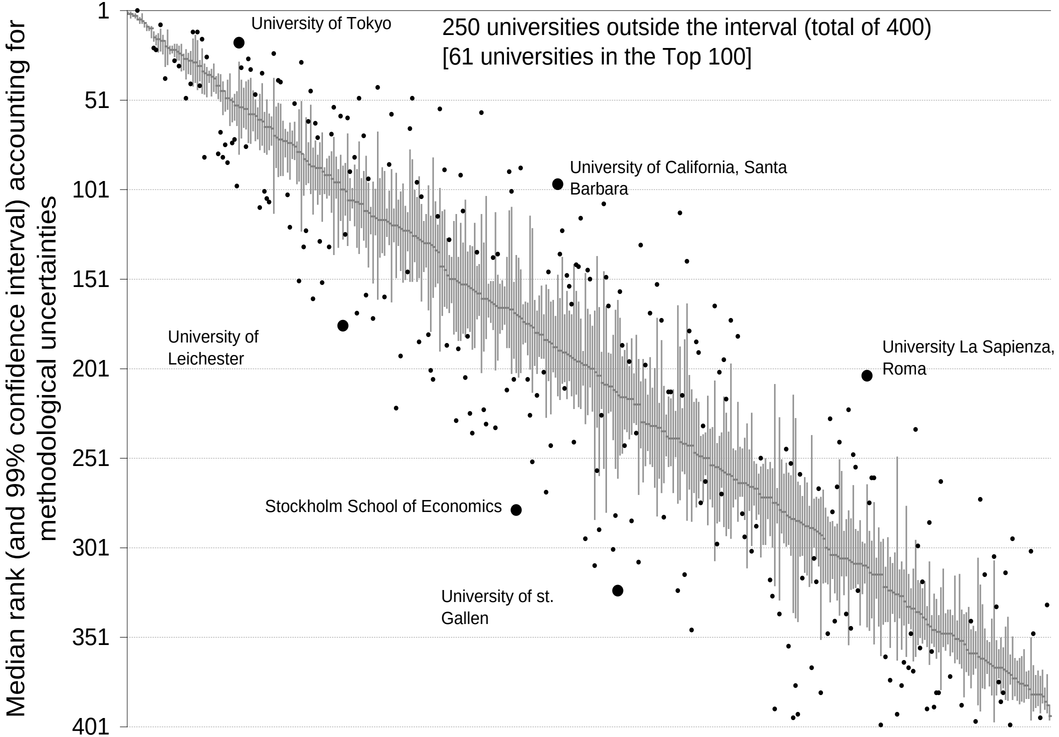
	1-5	6-10	11-15	16-20	21-25	26-30	31-35	36-40	41-45	46-50	51-55	56-60	61-65	66-70	71-75	76-80	81-85	86-90	91-95	96-100		
HARVARD University	44	56																			1	USA
YALE University	40	49	11																		2	USA
University of CAMBRIDGE	99																				3	UK
University of OXFORD	93	7																			4	UK
CALIFORNIA Institute of Technology	46	50																			5	USA
IMPERIAL College London	74	24																			6	UK
UCL (University College London)	73	23																			7	UK
University of CHICAGO		80	19																		8	USA
MASSACHUSETTS Institute of Technology	14	13	17	16	11	11	7														9	USA
COLUMBIA University	6	13	17	11	10	7	10	14													10	USA
University of PENNSYLVANIA		37	56	6																	11	USA
PRINCETON University	6	59	27	9																	12	USA
DUKE University			27	11	9		7		10	6	9		6								13	USA
JOHNS HOPKINS University			20	10	9	9	7		10		6	6	7			6					13	USA
CORNELL University		6	24	11		7	6	7	9	9		7									15	USA
AUSTRALIAN National University	10	30	29	31																	16	Australia
STANFORD University			10	14	7	10	9	10	6	6		7									17	USA
University of MICHIGAN			6	27	17	9	10	7	14	6											18	USA
University of TOKYO				16	7	13	7										6		6		19	Japan
MCGILL University			7	19	41	13	9	7													20	Canada

- Impact of uncertainties on the university ranks is even more apparent.
- M.I.T.: ranked 9th, but confirmed only in 13% of simulations (plausible range [4, 35])
- Very high volatility also for universities ranked 10th-20th position, e.g., Duke Univ, John Hopkins Univ, Cornell Univ.

Uncertainty analysis – ARWU results



Uncertainty analysis – THE results



European Commission

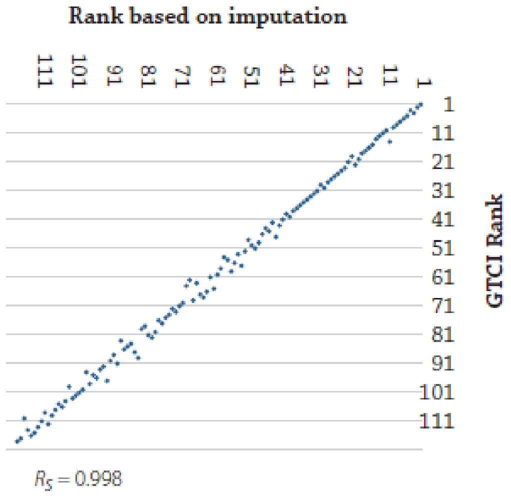
The Global Talent Competitiveness Index 2017 (GTCI)
Talent and Technology
TECHNOLOGY & WORK: DISRUPTION & CREATION
THE ADVANCE OF TECHNOLOGY IS
DISRUPTING
THE WORLD OF WORK
200 MILLION PEOPLE
ARE UNEMPLOYED IN THE WORLD AND
MANY JOBS ARE AT RISK
BUT TECHNOLOGY WON'T KILL WORK:
IT STIMULATES GROWTH AND CREATES NEW JOBS
THE NEW NATURE OF WORK
HIGH CONNECTEDNESS:
Collaboration and co-creation
WORK LIFE BLEND
THE JOB FOR LIFE NO LONGER EXISTS:
Multi-career is the norm
BEYOND AUTOMATION
NEW JOBS REQUIRE
NEW SKILLS & ATTITUDES
FLEXIBILITY
IS KEY
30%
of
US & EUROPEAN
workers are
FREE AGENTS
THE GTCI RANKS COUNTRIES BY THEIR ABILITY TO GROW, ATTRACT AND RETAIN TALENT

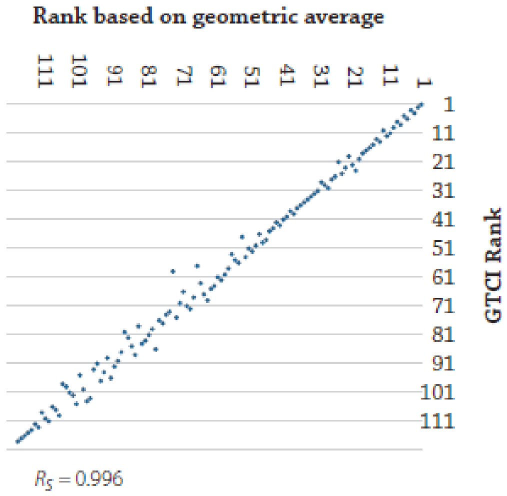
Uncertainty analysis for the GTCI 2017: Weights, missing data, aggregation, and normalisation

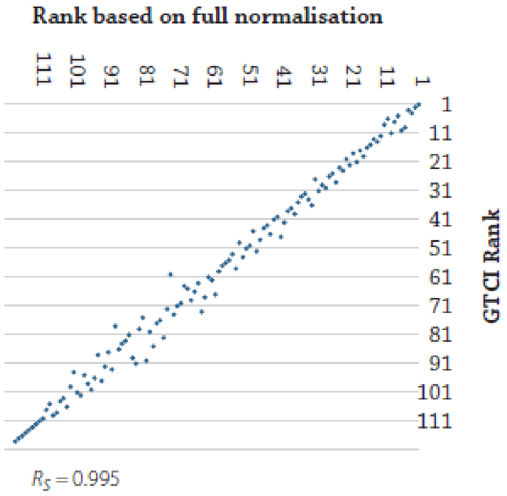
		REFERENCE	ALTERNATIVE
I. Uncertainty in the treatment of missing values		**No estimation of missing data**	**Expectation Maximisation (EM)**
II. Uncertainty in the aggregation formula at pillar level		**Arithmetic average**	**Geometric average**
III. Uncertainty in the method of normalisation		**Partial normalisation**	**Full normalisation**
IV. Uncertainty in the weights			
GTCI sub-index	Pillar	Reference value for the weight (within the sub-index)	Distribution assigned for robustness analysis (within the sub-index)
Input	Enable	0.25	U[0.15,0.35]
	Attract	0.25	U[0.15,0.35]
	Grow	0.25	U[0.15,0.35]
	Retain	0.25	U[0.15,0.35]
Output	Vocational and Technical Skills	0.50	U[0.40,0.60]
	Global Knowledge Skills	0.50	U[0.40,0.60]

Sensitivity Analysis

Which assumptions are driving the uncertainty in the rankings?



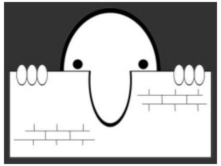




We can also go further and delve into the sensitivity analysis literature (variance-based SA, moment-independent measures, etc).

European Commission

A brief glimpse of SA...



If we want to go a step further we can calculate formal sensitivity measures telling us, e.g:

"Of the assumptions tested, the uncertainty in the choice of normalisation method causes 40% of the uncertainty in rankings"

Let A_1, A_2, A_3,... be the assumptions used in the composite indicator construction, which result in a set of scores:

$$\boxed{\text{scores} = \text{CI}(A_1, A_2, A_3, A_4, \ldots)}$$

The uncertainty in the *scores*, as a result of uncertainty in the A_i, can be captured by var(scores). Sensitivity analysis can decompose this variance into portions attributable to each assumption.

$$\mathrm{var}(\mathrm{scores}) = \sum_{i=1}^{k} V_i + \sum_{i}\sum_{j<i} V_{i,j} + \cdots + V_{1,2,\ldots,k}$$

$$V_i = \mathrm{var}_{A_i}[\mathrm{E}_{A_{\sim i}}(\mathrm{scores}|A_i)]$$

$$V_{i,j} = \mathrm{var}_{A_i,A_j}\left[\mathrm{E}_{A_{\sim i,j}}\left(\mathrm{scores}\middle|A_i, A_j\right)\right] - \mathrm{var}_{A_i}[\mathrm{E}_{A_{\sim i}}(\mathrm{scores}|A_i)] - \mathrm{var}_{A_j}[\mathrm{E}_{A_{\sim j}}\left(\mathrm{scores}\middle|A_j\right)]$$

Requires a specific design.

Very informative but perhaps hard to communicate...

Saisana M., Saltelli A., Tarantola S., 2005, Uncertainty and sensitivity analysis techniques as tools for the analysis and validation of composite indicators. *J Royal Statistical Society A* **168(2)**, 307-323

If you want to know more…

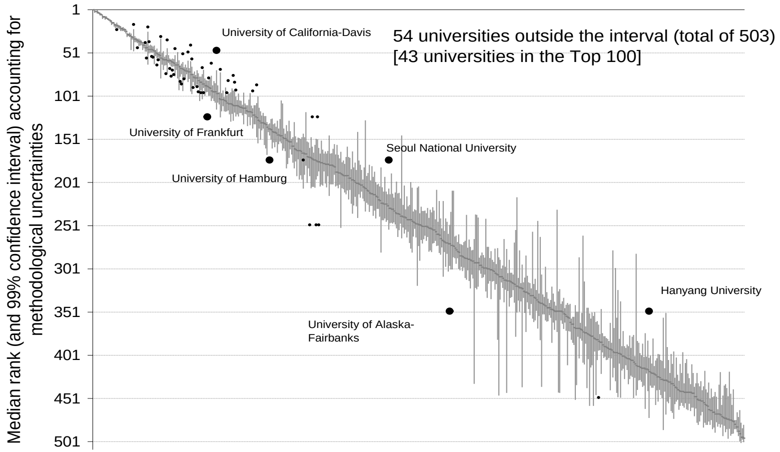
Summer school on sensitivity analysis

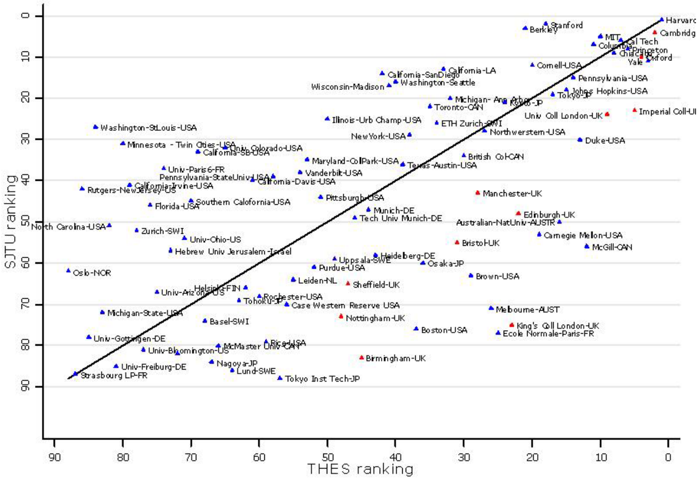
June 11-15, Ispra (here)

Further details to be announced
Contact me if interested

Uncertainty and validation

We cannot explore the full uncertainty of a composite indicator: we can only explore some of the assumptions.





“Given the assumptions that were tested, the outcome of the composite indicator is shown to vary in the following way...” [lower bound on uncertainty]

In general it is not possible to *validate* a composite indicator (build compelling evidence that it is an effective model).
(but, the same problems apply to some extent to any model)

From irresponsible musings to serious efforts

Rankings range from irresponsible musings by self-appointed experts and money-making schemes by commercial organizations to, at their best, serious efforts by academic or research organizations. (Aitbach, 2015)

Composite indicators

- Are very widely used
- Fill a demand for which there is no other alternative

Therefore we should use available tools to **increase robustness and credibility**:

1. Transparency—detailed description of methodology, data sources, assumptions
2. Statistical soundness—analysis of correlations, data structure, effects of weights, etc.
3. Uncertainty and sensitivity analysis—check effect of alternative but plausible assumptions. Honestly acknowledge uncertainty.

'rankings are here to stay, and it is therefore worth the time and effort to <u>get them right</u>'

(Alan Gilbert, Nature News, 2007)

'composite indicators are here to stay, and it is therefore worth the time and effort to develop them responsibly and use them sensibly'

(JRC, Composite Indicators Research Group, 2017)

References

Aitbach, P. (2015). The dilemmas of ranking. *International Higher Education*, (42)

Becker, W., Saisana, M., Paruolo, P., & Vandecasteele, I. (2017). Weights and importance in composite indicators: Closing the gap. *Ecological Indicators*, *80*, 12-22.

Booysen, F. (2002). An overview and evaluation of composite indices of development. *Social indicators research*, *59*(2), 115-151

Saisana, M., Saltelli, A., & Tarantola, S. (2005). Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *168*(2), 307-323.

Saisana, M., d'Hombres, B., & Saltelli, A. (2011). Rickety numbers: Volatility of university rankings and policy implications. *Research policy*, *40*(1), 165-177.

THANK YOU

Welcome to email us at: jrc-coin@ec.europa.eu

COIN in the EU Science Hub
https://ec.europa.eu/jrc/en/coin

COIN tools are available at:
https://composite-indicators.jrc.ec.europa.eu/

European Commission