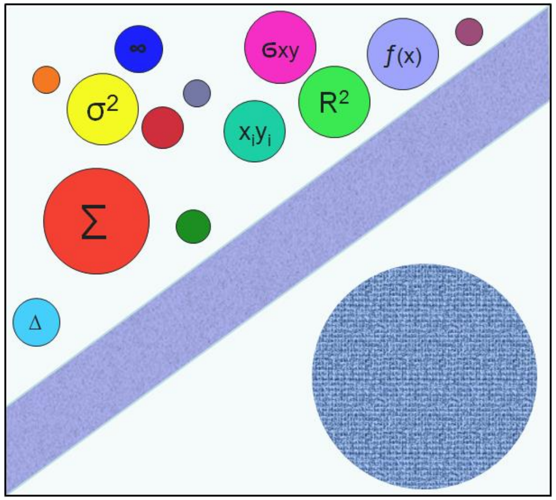
On a Generalized Non-Compensatory Composite Index for Measuring Socio-economic Phenomena

Matteo Mazziotta and Adriano Pareto
Italian National Institute of Statistics



Conference
"Dealing with complexity in society:
from plurality of data to synthetic indicators"
Padua, 17-18 Sept 2015

Contents

Introduction

Social and economic phenomena, such as development, poverty, quality of life, are difficult to measure and evaluate since they are characterized by a multiplicity of aspects or dimensions.

The complex and multidimensional nature of these phenomena requires the definition of intermediate objectives whose achievement can be observed and measured by individual indicators.

A *composite index* is a mathematical combination (or aggregation) of a set of individual indicators that represent the different dimensions of the phenomenon to be measured.

Summarizing complex and multidimensional phenomena into single numbers involves several theoretical, methodological and empirical problems.

Constructing a composite index

Defining the phenomenon to be measured. The definition of the concept should give a clear sense of what is being measured by the composite index.

Selecting a group of individual indicators. Indicators should be selected according to their relevance, analytical soundness, time-liness, accessibility and so on.

Normalizing the individual indicators. This step aims to make the indicators comparable. Normalization is required before any data aggregation as the indicators in a data set often have different measurement units.

Aggregating the normalized indicators. It is the combination of all the components to form one or more composite indices (mathematical functions).

A non-compensatory composite index

Basic idea

The MPI (Mazziotta-Pareto Index) is a formative composite index for summarizing a set of indicators that are assumed to be *non-substitutable*, i.e., all components must be *balanced*.

It is based on a non-linear function which, starting from the arithmetic mean, introduces a penalty for the units with unbalanced values.

Requirements

- Independence from the unit of measurement and the variability of the indicators
- Independence from an “ideal unit”, since it is arbitrary, non-univocal and can vary with time
- Simplicity of computation
- Ease of interpretation

Method and formula

Given the matrix $\mathbf{X} = \{x_{ij}\}$ with n rows (units) and m columns (indicators), we calculate the normalized matrix $\mathbf{Z} = \{z_{ij}\}$ where the indicator j has mean 100 and standard deviation 10.

Denoting with M_i, S_i, cv_i, respectively, the mean, the standard deviation and the coefficient of variation of z_{ij} $\{j = 1,\ldots, m\}$, the generalized form of the MPI is given by:

$$\mathrm{MPI}_i^{+/-} = \mathrm{M}_i \pm \mathrm{S}_i \, \mathrm{cv}_i$$

where the sign ± depends on the kind of phenomenon to be measured (in case of development, the MPI^- is used).

Therefore, the MPI is characterized by the combination of a “mean effect” (M_i) and a “penalty effect” ($S_i \, cv_i$).

Some properties

(i) The MPI^+ and the MPI^- of unit *i* are *reflexive*, i.e., if $z_{ij} = z_i$ ($j = 1,\ldots, m$), that is $S_i = 0$, then:

$$MPI_i^+ = MPI_i^- = z_i$$

(ii) The MPI^+ of unit *i* is greater or equal than the MPI- of the same unit, that is:

$$MPI_i^+ \geq MPI_i^-$$

In particular, $MPI_i^+ = MPI_i^-$ iff $S_i = 0$

(iii) The MPI^+ and the MPI^- of the unit *i* are linked by the relation:

$$MPI_i^- = 2M_i - MPI_i^+ \qquad \text{or} \qquad \frac{MPI_i^- + MPI_i^+}{2} = M_i$$

Some properties

(iv) Given two units i and h ($i \neq h$), with $M_i = M_h$, we have:

$$\mathrm{MPI}_i^- > \mathrm{MPI}_h^- \quad \text{iff} \quad S_h > S_i$$
$$\mathrm{MPI}_i^+ > \mathrm{MPI}_h^+ \quad \text{iff} \quad S_i > S_h$$

(v) Given two units i and h ($i \neq h$), with $M_i > M_h$, we have:

$$\mathrm{MPI}_i^- > \mathrm{MPI}_h^- \quad \text{iff} \quad M_i - M_h > S_i\, cv_i - S_h\, cv_h$$
$$\mathrm{MPI}_i^+ > \mathrm{MPI}_h^+ \quad \text{iff} \quad M_i - M_h > S_h\, cv_h - S_i\, cv_i$$

(vi) Let $r(x_j, x_k)$ be the linear correlation coefficient between the indicators j and k; if $r(x_j, x_k) = 1$, for each j and k ($j \neq k$), then:

$$\mathrm{MPI}_i^+ = \mathrm{MPI}_i^- = M_i$$

Istat

Some properties

$r(X_1, X_2)=1$ (maximum positive correlation)

Unit	Original indicators		Normalized indicators		Mean	Std. dev.	MPI+	MPI-
	X1	X2	Z1	Z2				
A	11	100	114,1	114,1	114,1	0,0	114,1	114,1
B	9	80	107,1	107,1	107,1	0,0	107,1	107,1
C	7	60	100,0	100,0	100,0	0,0	100,0	100,0
D	5	40	92,9	92,9	92,9	0,0	92,9	92,9
E	3	20	85,9	85,9	85,9	0,0	85,9	85,9
Mean	**7**	**60**	**100**	**100**				
Std. dev.	**2,8**	**28,3**	**10**	**10**				

With maximum positive correlation, all the units have standard deviation equal to 0 and the MPI depends exclusively on the mean.
Thus, the MPI ranks the units according to the *mean* level.

Some properties

$r(X_1, X_2)$=-1 (maximum negative correlation)

Unit	Original indicators		Normalized indicators		Mean	Std. dev.	MPI+	MPI-
	X1	X2	Z1	Z2				
A	3	100	85,9	114,1	100,0	14,1	102,0	98,0
B	5	80	92,9	107,1	100,0	7,1	100,5	99,5
C	7	60	100,0	100,0	100,0	0,0	100,0	100,0
D	9	40	107,1	92,9	100,0	7,1	100,5	99,5
E	11	20	114,1	85,9	100,0	14,1	102,0	98,0
Mean	**7**	**60**	**100**	**100**				
Std. dev.	**2,8**	**28,3**	**10**	**10**				

With maximum negative correlation, all the units have mean equal to 100 and the MPI depends exclusively on the standard deviation.
Thus, the MPI ranks the units according to the *variability* level.

Some properties

$r(X_1, X_2)=0$ (zero correlation)

Unit	Original indicators		Normalized indicators		Mean	Std. dev.	MPI+	MPI-
	X1	X2	Z1	Z2				
A	11	100	88,4	114,1	101,3	12,9	102,9	99,6
B	16	80	110,7	107,1	108,9	1,8	108,9	108,9
C	14	60	101,8	100,0	100,9	0,9	100,9	100,9
D	16	40	110,7	92,9	101,8	8,9	102,6	101,0
E	11	20	88,4	85,9	87,1	1,3	87,2	87,1
Mean	**13,6**	**60**	**100**	**100**				
Std. dev.	**2,2**	**28,3**	**10**	**10**				

In the intermediate case, the MPI depends on the mean and the standard deviation.
Thus, the MPI is a combination of both the *mean* and the *variability* level.

A variant for spatio-temporal comparisons

Given the matrix **X** = $\{x_{ij}\}$ with *n* rows (units) and *m* columns (indicators), we calculate the normalized matrix **R** = $\{r_{ij}\}$ as follow:

$$r_{ij} = \frac{(x_{ij} - \text{Min}_{x_j})}{(\text{Max}_{x_j} - \text{Min}_{x_j})} 60 + 70$$

where Min_{x_j} and Max_{x_j} are the ‘goalposts’ for the indicator *j*.

Denoting with M_i, S_i, cv_i, respectively, the mean, the standard deviation and the coefficient of variation of r_{ij} $\{j = 1,\ldots, m\}$, the generalized form of the AMPI (Adjusted MPI) is given by:

$$\text{AMPI}_i^{+/-} = M_i \pm S_i \, cv_i$$

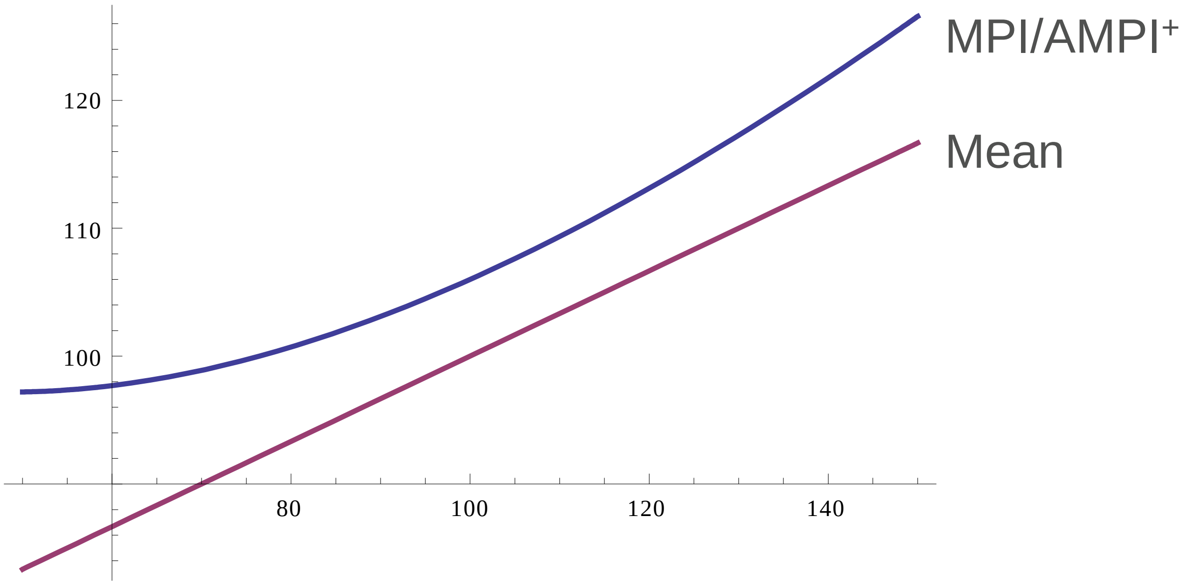
where the sign ± depends on the kind of phenomenon to be measured (in case of development, the AMPI^- is used).

Theoretical aspects

The positive penalty index

The MPI/AMPI$^+$ is a *convex* function of z_{ik}/r_{ik} (k = 1,…, m) and may be considered monotonic increasing in the range 70-130.

Example:

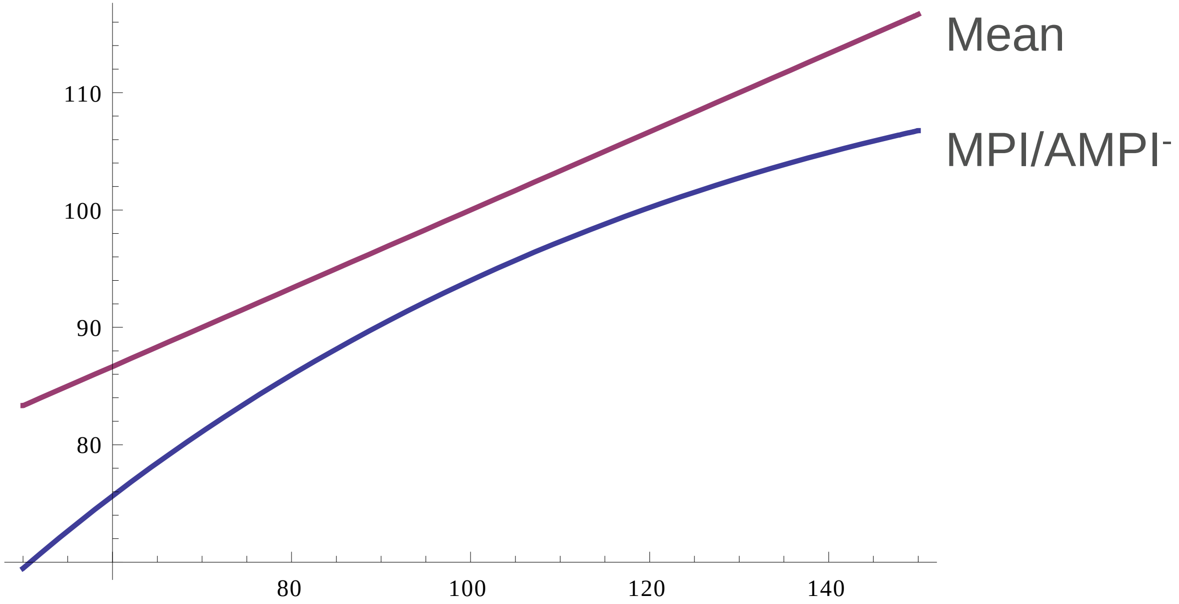


Theoretical aspects

The negative penalty index

The MPI/AMPI$^-$ is a *concave* function of z_{ik}/r_{ik} (k = 1,…, m) and may be considered monotonic increasing in the range 70-130.

Example:



An application to well-being indicators

List of indicators of well-being and definitions

Indicator	Definition
Life expectancy	It is the standard measure of the length of people's life. Life-expectancy measures how long on average people could expect to live based on the age specific mortality rates currently prevailing. Life-expectancy can be computed at birth and at various ages
Educational attainment	It profiles the education of the adult population as captured through formal educational qualifications. Educational attainment is measured as the percentage of the adult population (15 to 64 years of age) holding at least an upper secondary degree, as defined by the OECD-ISCED classification
Employment rate	It is the share of the working age population (people aged from 15 to 64 in most OECD countries) who are currently employed in a paid job. Employed persons are those aged 15 and over who declare having worked in gainful employment for at least one hour in the previous week, following the standard ILO definition
Household disposable income	It includes income from work, property, imputed rents attributed to home owners and social benefits in cash, net of direct taxes and social security contributions paid by households; it also includes the social transfers in kind, such as education and health care, that households receive from governments. Income is measured net of the depreciation of capital goods that households use in production

Source: www.oecdbetterlifeindex.org

An application to well-being indicators

Individual indicators of well-being - Original values (Years 2011, 2014)

Country	2011				2014			
	Life expectancy	Educational attainment	Employment rate	Household disposable income	Life expectancy	Educational attainment	Employment rate	Household disposable income
Australia	81.5	69.7	72.3	27,039	82.0	74.0	72.0	31,197
Austria	80.5	81.0	71.7	27,670	81.1	82.0	73.0	29,256
Belgium	79.8	69.6	62.0	26,008	80.5	71.0	62.0	27,811
Canada	80.7	87.1	71.7	27,015	81.0	89.0	72.0	30,212
Chile	77.8	68.0	59.3	8,712	78.3	72.0	62.0	13,762
Czech Republ	77.3	90.9	65.0	16,690	78.0	92.0	67.0	17,262
Denmark	78.8	74.6	73.4	22,929	79.9	77.0	73.0	25,172
Estonia	73.9	88.5	61.0	13,486	76.3	89.0	67.0	14,382
Finland	79.9	81.1	68.1	24,246	80.6	84.0	70.0	26,904
France	81.0	70.0	64.0	27,508	82.2	72.0	64.0	29,322
Germany	80.2	85.3	71.1	27,665	80.8	86.0	73.0	30,721
Greece	80.0	61.1	59.6	21,499	80.8	67.0	51.0	19,095
Hungary	73.8	79.7	55.4	13,858	75.0	82.0	57.0	15,240
Ireland	79.9	69.5	60.0	24,313	80.6	73.0	59.0	23,721
Italy	81.5	53.3	56.9	24,383	82.7	56.0	58.0	24,724
Japan	82.7	87.0	70.1	23,210	82.7	93.0	71.0	25,066
Korea	79.9	79.1	63.3	16,254	81.1	81.0	64.0	18,035
Mexico	75.1	33.6	60.4	12,182	74.4	36.0	61.0	12,850
Netherlands	80.2	73.3	74.7	25,977	81.3	72.0	75.0	25,697
New Zealand	80.4	72.1	72.3	18,819	81.2	74.0	72.0	21,773
Norway	80.6	80.7	75.3	29,366	81.4	82.0	76.0	32,093
Poland	75.6	87.1	59.3	13,811	76.9	89.0	60.0	16,234
Portugal	79.3	28.2	65.6	18,540	80.8	35.0	62.0	18,806
Slovak Repub	74.8	89.9	58.8	15,490	76.1	91.0	60.0	17,228
Slovenia	78.8	82.0	66.2	19,890	80.1	84.0	64.0	19,692
Spain	81.2	51.2	58.6	22,972	82.4	54.0	56.0	22,799
Sweden	81.2	85.0	72.7	26,543	81.9	87.0	74.0	27,546
Switzerland	82.2	86.8	78.6	27,542	82.8	86.0	79.0	30,745
United Kingdo	79.7	69.6	69.5	27,208	81.1	77.0	71.0	25,828
United States	77.9	88.7	66.7	37,685	78.7	89.0	67.0	39,531
Average	**79.2**	**74.1**	**66.1**	**22,284**	**80.1**	**76.5**	**66.4**	**23,757**

Source: www.oecdbetterlifeindex.org

An application to well-being indicators

MPI of well-being (Years 2011, 2014 and variation)

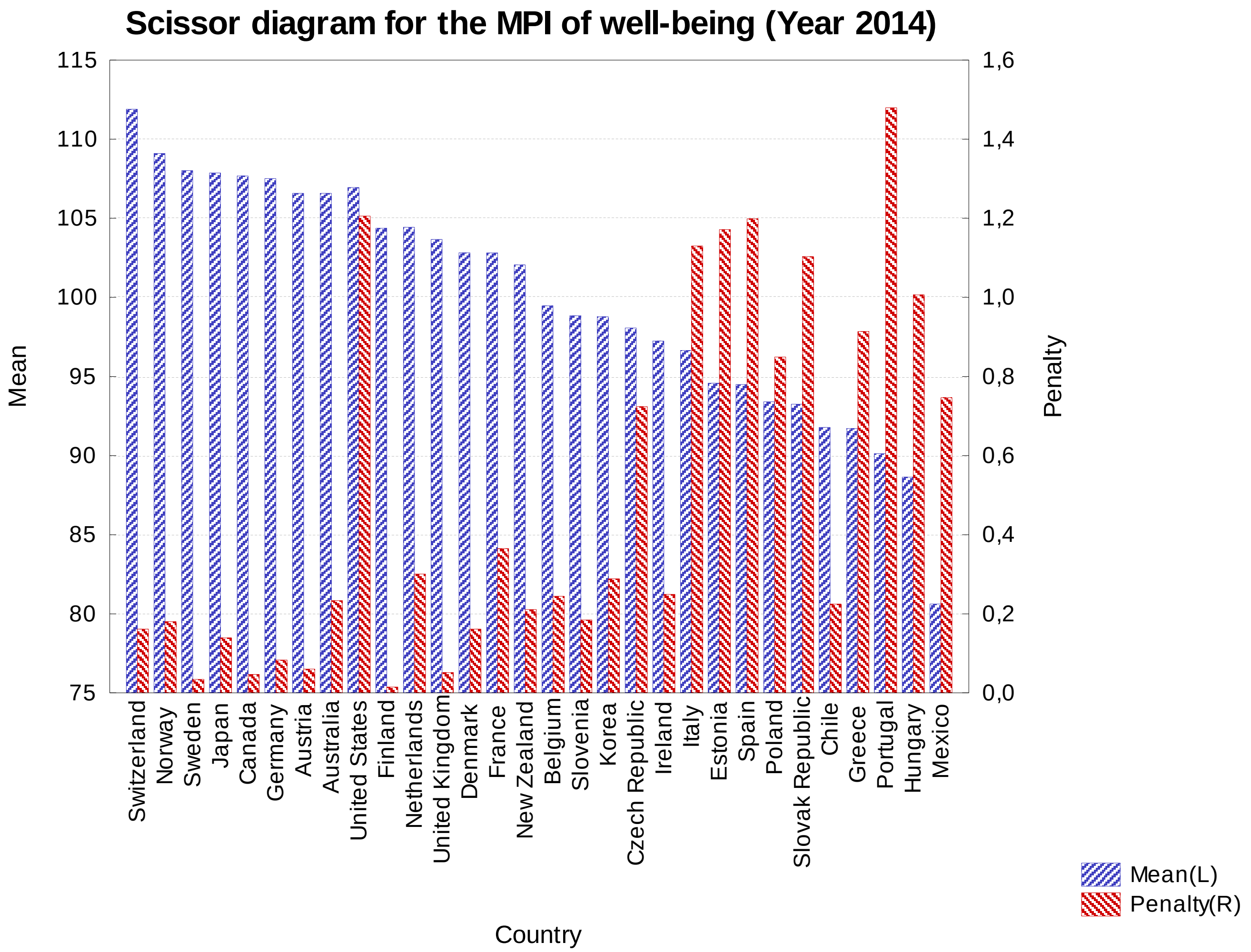
Country	2011		2014		Variation	
	Value	Rank	Value	Rank	Value	Rank
Australia	105.80	9	106.35	8	0.55	9
Austria	106.83	7	106.52	7	-0.30	21
Belgium	99.50	17	99.22	16	-0.28	20
Canada	107.76	4	107.63	5	-0.13	18
Chile	88.89	28	91.56	26	2.67	2
Czech Republic	97.40	20	97.35	19	-0.05	17
Denmark	102.55	13	102.67	13	0.11	15
Estonia	89.73	26	93.41	22	3.68	1
Finland	103.44	11	104.36	10	0.92	6
France	102.15	14	102.45	14	0.30	12
Germany	106.96	6	107.43	6	0.46	11
Greece	95.47	21	90.80	27	-4.67	30
Hungary	86.48	29	87.63	29	1.15	4
Ireland	98.04	18	96.99	20	-1.05	25
Italy	95.04	22	95.51	21	0.47	10
Japan	107.51	5	107.73	4	0.22	13
Korea	97.73	19	98.49	18	0.76	7
Mexico	82.34	30	79.87	30	-2.48	29
Netherlands	105.54	10	104.13	11	-1.41	28
New Zealand	101.68	15	101.84	15	0.17	14
Norway	108.85	2	108.93	2	0.07	16
Poland	91.23	24	92.55	24	1.32	3
Portugal	89.32	27	88.64	28	-0.68	24
Slovak Republic	91.02	25	92.15	25	1.12	5
Slovenia	99.82	16	98.66	17	-1.17	26
Spain	94.70	23	93.30	23	-1.41	27
Sweden	108.18	3	107.99	3	-0.19	19
Switzerland	112.07	1	111.74	1	-0.34	22
United Kingdom	102.94	12	103.62	12	0.67	8
United States	106.18	8	105.75	9	-0.43	23
Average	**100.00**		**100.00**			

An application to well-being indicators

AMPI of well-being (Years 2011, 2014 and variation)

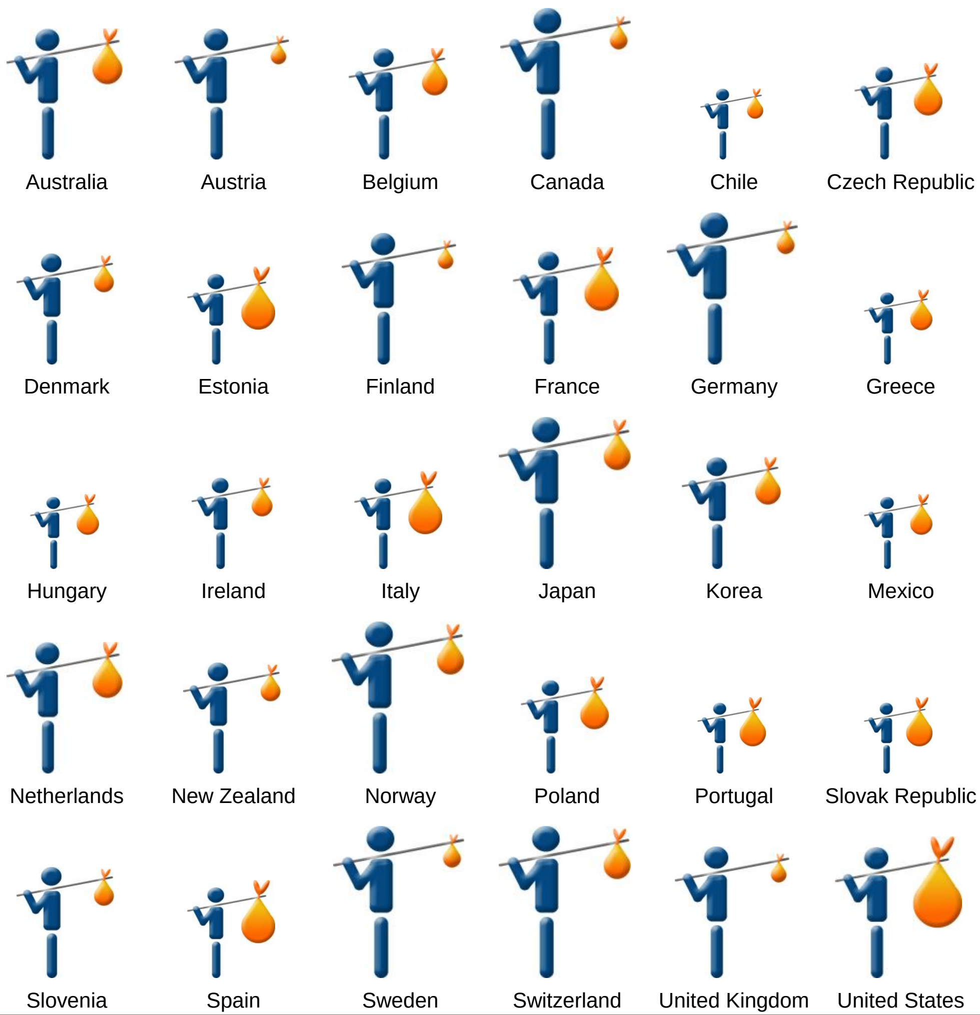
Country	2011		2014		Variation	
	Value	Rank	Value	Rank	Value	Rank
Australia	107.91	8	111.63	8	3.72	8
Austria	109.34	7	111.99	7	2.64	21
Belgium	99.14	17	101.27	16	2.13	24
Canada	110.76	5	113.43	5	2.67	19
Chile	85.27	26	91.33	25	6.06	2
Czech Republic	96.10	20	99.06	19	2.96	14
Denmark	103.17	13	106.71	13	3.54	10
Estonia	84.00	27	93.35	22	9.35	1
Finland	104.80	11	108.94	10	4.14	7
France	102.86	14	105.79	14	2.93	15
Germany	109.51	6	113.15	6	3.63	9
Greece	93.74	21	88.77	27	-4.97	30
Hungary	79.74	29	84.53	29	4.79	3
Ireland	97.23	19	98.13	20	0.90	27
Italy	92.67	22	95.57	21	2.90	17
Japan	110.85	4	113.76	4	2.91	16
Korea	97.37	18	100.86	18	3.49	11
Mexico	74.61	30	74.52	30	-0.09	29
Netherlands	107.41	9	108.84	11	1.43	25
New Zealand	102.55	15	105.77	15	3.22	13
Norway	111.98	2	115.30	2	3.32	12
Poland	87.12	24	91.88	24	4.76	5
Portugal	83.89	28	86.56	28	2.66	20
Slovak Republic	86.20	25	90.99	26	4.79	4
Slovenia	99.83	16	100.99	17	1.16	26
Spain	92.28	23	92.37	23	0.09	28
Sweden	111.41	3	114.16	3	2.75	18
Switzerland	116.79	1	119.41	1	2.62	22
United Kingdom	103.73	12	108.03	12	4.30	6
United States	107.10	10	109.66	9	2.56	23
Average	**100.00**		**102.86**		**2.86**	

An application to well-being indicators



An application to well-being indicators

Traveller Icon plots of the MPI of well-being (Year 2014)



Conclusions

In this work, a generalized non-compensatory composite index (MPI), and its variant for spatio-temporal comparisons (AMPI), were considered and their main properties were examined

The main difference between MPI and AMPI is the normalization method:

- The MPI is based on a standardization of the individual indicators and measures only relative differences with respect to the mean
- The AMPI is based on a re-scaling of the individual indicators and measures absolute differences with respect to prefixed goalposts

The MPI is the best solution for a 'static' analysis (e.g., a single-year analysis), whereas the AMPI is the best solution for a 'dynamic' analysis (e.g., a multi-year analysis)

References

De Muro, P., Mazziotta, M., & Pareto, A. (2011). Composite Indices of Development and Poverty: An Application to MDGs. *Soc. Indic. Res*., 104, 1-18.

Mazziotta M., Pareto A. (2015). On a Generalized Non-compensatory Composite Index for Measuring Socio-economic Phenomena. *Soc. Indic. Res*, doi: DOI 10.1007/s11205-015-0998-2. Springer.

OECD (2008). *Handbook on Constructing Composite Indicators. Methodology and user guide*. Paris: OECD Publications.

Salzman, J. (2003). *Methodological Choices Encountered in the Construction of Composite Indices of Economic and Social Well-Being*. Technical Report, Center for the Study of Living Standards, Ottawa.