Is Happiness generated by substitutable inputs? Preliminary techniques and implications for both policy and indicator makers

Demetrio Miloslavo Bova[1]

Abstract *This dissertation enquires whenever the factors affecting happiness are substitutes by introducing a pilot study with a new statistical technique. Perfect substitutability would entail important differences in the way we define policies and indicators. Indeed, policies may focus on each aspect separately and composite indicators could be simple weighted average. The model developed is applied to the world happiness report data.*

Keywords: Substitutability, Happiness, S-Weighted average, Complementarity, Leaf graph.

Gruppo tematico: 19. Costruzione indicatori e loro sintesi

1 Introduction: the relevance of the issue

The relevance of wellbeing and happiness as the main parameter to measure the progress of society is more and more established in research as well as in political life [1,2]. Maybe the main challenge was, and perhaps still is, to avoid that the only indicator of GDP can be used as a unique measure although its link with happiness is non neglectable [3]. The attempt to go beyond the GDP requires the presence of models able to address happiness including also other factors. Here I focus on the possibility that substitute inputs explain happiness.

We often use models where the utility function includes complementary 'inputs' and indicators based on weighted averages. In case of complementary, the marginal benefit of an input depends on its relative size while in case of substitutability it depends on its size only. In terms of indicators, the first case requires a complex ponderation of the inputs, in the latter, each input/indicator can be treated independently.

I will briefly introduce a model to test perfect substitutability, to check if a weighted average may simulate it, a fast application on the world happiness report data, and the consequent comments.

[1] Demetrio Miloslavo Bova, University of Warsaw, dbova@coin.wne.uw.edu.pl

2 Model

In this section I will introduce a model to test the perfect substitutability of inputs and the possibility that a weighted average can simulate such a relationship.

2.1 Substitutability

An input output relationship is defined perfect substitutability if, given n inputs (I), each of them concurs independently to the generation of the output (Out). The relationship can be described as

$$\sum_{i=1}^{i=n} \mathrm{a_i}\, I_i \;\; = Out + \varepsilon, \;\; \forall\ \mathrm{i} \in (1, \mathrm{n})\ \mathrm{a_i} > 0 \tag{1}$$

Where a is the slope and ε is the error. Let me focus on rescaled variables (r) according to the following standard formula.[1]

$$I_r = \frac{I - \min I}{\max I - \min I} \tag{2}$$

The graphical representation for two substitute inputs is the following.
A rescaled input is expected to have a positive slope and an intercept higher than zero if regressed with respect to the output (left graphs). The errors size is due to the impact of the other inputs such that the intercept of the previous regression is positive and informs about the average effect of the other inputs.

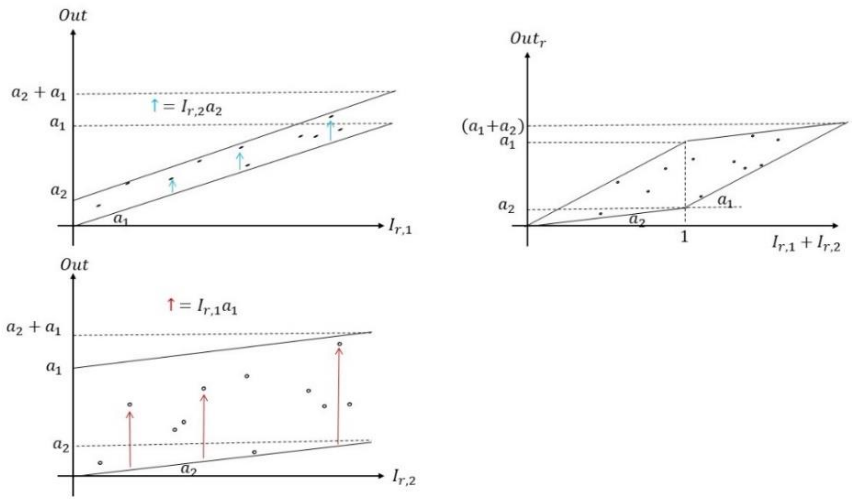

Figure 1: Substitute rescaled inputs and rescaled output.

[1] Such a formula hides a potential deformation if deal with indicators, as the GDP, that have no maximum or minimum definable ex-ante.

Moreover, if all the inputs are substitutes, considering the output rescaled, then the sum of the rescaled inputs slopes times their maximum values must be equal to one since one is the maximum of the rescaled output (right graph).

Not all the inputs have the same impact on the output. When we consider the sum of the rescaled inputs it may happen that the input used are only the lower or the higher in terms of impact. Hence, for low values of the sum of the rescaled inputs we have a range of acceptable values between the lower and the higher. When the sum of rescaled input increases and goes over 1/n, then is no more possible that the inputs used are only those the higher and the lower productivity, hence, we should consider the second higher and lower as well. This consideration repeats until we reach the maximum sum of rescaled inputs where the slope becomes necessarily equal to one. The graph encompassing the possible observations in the plan with rescaled sum of the inputs and rescaled output will be labelled Leaf graph and the upper and lower border are defined as follows.

$$Upper\ Leaf_j = \sum_{i=1}^{i=j}(a_i)\,,\ Lower\ Leaf_j = \sum_{i=j}^{i=n}(a_i)\ ,\ \ where\ \forall(i<j)\ a_i \geq a_j \qquad (3)$$

If the elements on the plan (average sum of rescaled input, rescaled output) are inside the leaf, and the previous conditions are met, then we satisfy some necessary (although not sufficient) conditions of substitutability. In this dissertation I will not enquire further aspects although they may be important.

For simplicity, I consider the sum pf input rescaled by the number of variables (n) (that is the average of input rescaled) to obtain a graph included in the square delimitated by the origin and the point (1,1).

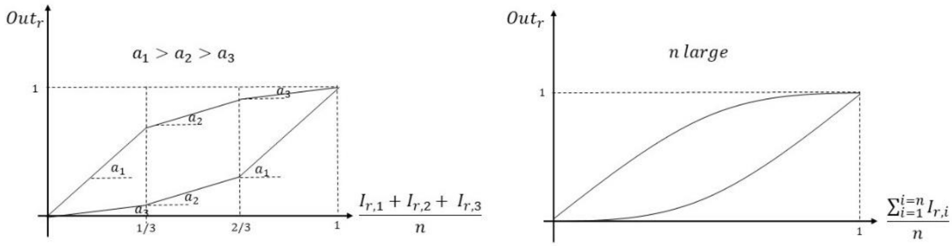

Figure 2: Leaf graph (upper) with three variables (lower) with n variables.

2.2 Test

To introduce the tests let me consider the plans $(I_{r,i}, Out)$ and $(\sum_{i=1}^{i=n} I_{r,i}\,, Out)$ and, for space reasons, let me rely on both the previous graphical considerations.

- Plan $(\mathbf{I_{r,i}}, \mathbf{Out})$

The substitutability can be tested by a set of regressions for each rescaled input versus the rescaled output where the hypothesises are: 1) the intercept is nonnegative; 2) the slope is positive and lower than one.

- Plan $(\sum_{i=1}^{i=n} \mathbf{I_{r,i}}\,, \mathbf{Out})$

The substitutability can be tested by a regression on the rescaled input sum versus the rescaled output and through the leaf, the hypothesises are: 1) the slope is equal to 1; 2) the intercept is equal to 0; 3) The elements are encompassed by the leaf.

2.3 Centred leaf: test for S-weighted averages

Let me label ***S-weighted average*** the sum of weighted variables where each weight (w) is the corresponding slope (a) of that rescaled independent variable regressed for the rescaled dependent variable divided by the sum of the slopes.

$$S - weighted\ average = w_i I_{r,i} = \frac{\sum(a_i)}{\sum a_i} I_{r,i} \quad (4)$$

As said, whenever the sum of the slopes ($\sum a_i$) is higher than one then we do not have substitution. However, it is possible due to variables reciprocal interference or correlation that the S-weighted average obtains discrete results in the description of the phenomenon. In this case, the indicator made by S-weighted average generates only the illusion that each element can be considered separately as a perfect substitute. To check this 'illusion' we can correct the leaf as follows and compare it with the previous one.

$$Centred\ Upper\ Leaf_j = s\frac{(Upper\ Leaf_j)}{\bar{a}} - q,\ Centred\ Lower\ Leaf_j = s\frac{(Lower\ Leaf_j)}{\bar{a}} - q \quad (5)$$

Where 's' and 'q' are, respectively, the slope and the intercept of the regression between the sum of rescaled inputs and the rescaled output and $\bar{a}$ is the average a. Whenever the S-weighted average is encompassed in this leaf we know that it may also succeeds to describe the phenomenon but cannot be interpreted as sum of substitutes if the previous conditions are not satisfied as well.

3 Data and results

This section presents the data used to test the previous model, the results, and concludes with a comment.

3.1 data

The variables selected are a subset of the world happiness report data [4].

Life ladder	Min=0	Max=10	**Freedom to make life choices**	Min=0	Max=1
GDP Per capita Log	Min=0	Max= -	**Democratic quality**	Min= -	Max=-
Social support	Min=0	Max=1	**Positive affect**	Min=0	Max=1
Health (life expectancy at birth)	Min=0	Max= -	**Delivery quality**	Min= -	Max= -
Other information					
Sample size		629	**Data source**	[4] and related statistical appendix	

Note: The minimums and maximums refer to the values used to rescale; they coincide with the indicator minimum /maximum only when it is bounded. For the others, the min and the max were computed as the min of the max of the sample increased by a factor of 0.66.

Table 1: Indicators selected and sample information

3.2 Results: substitution

Indicator Rescaled	**Intercept**	**P value**	**Slope**	**P value**	**R squared**
GDP per capita (log)	-0,15378	<0,001	1,44572004	<0,001	0,550850
Social Support	0,049666	0,056	0,61758379	<0,001	0,386082
Health (life expectancy at birth)	-0,04524	0,079	1,20667656	<0,001	0,475009
Freedom to Make Life Choices	0,261168	<0,001	0,42032838	<0,001	0,322644
Positive Affect	0,172599	<0,001	0,54510214	<0,001	0,329709
Democratic Quality	0,31155	<0,001	0,43533193	<0,001	0,337974
Delivery Quality	0,340541	<0,001	0,474397	<0,001	0,449124
			Sum: 5,14		

Table 2 Linear regression on each rescaled indicator versus the rescaled Life Ladder

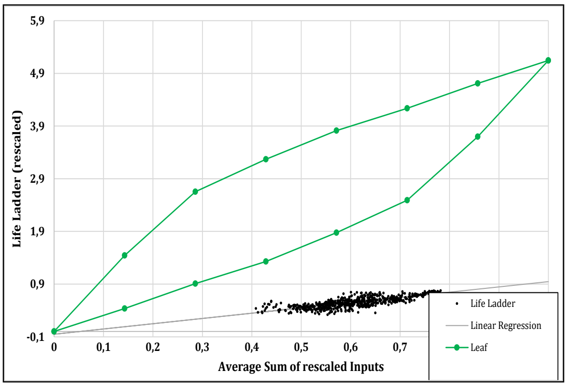

Figure 3: Leaf graph

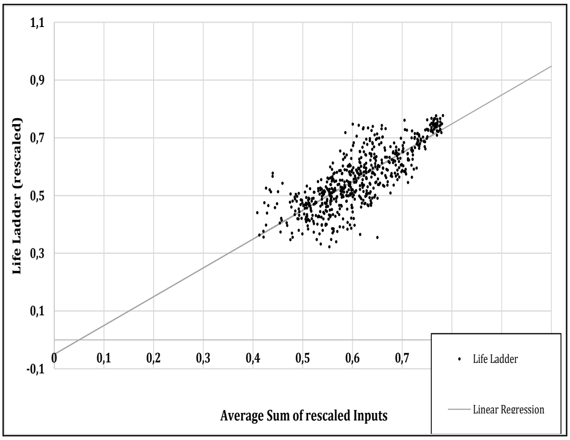

Figure 4: Life ladder rescaled Versus Average sum of rescaled inputs

$$Life\ Ladder_{0\to1} = \beta_0 + \beta_1(Sum\ of\ Rescaled\ Inputs)$$

	Coefficients	Standard Errors.	T-ratio	P-value
Constant	β_0= −0,0503713	0,0184230	−2,734	0,0064
Sum of Rescaled Inputs	β_1= 0,998465	0,029277	33,47	<0,0001
R-squared	0,641209		Akaike Criterion	−1688,029

Table 3: Results about rescaled Life ladder and Sum of Rescaled Inputs

3.3 Results: weighted average

	Log GDP per capita	Social support	Healthy life expectancy at birth	Freedom to make life choices	Positive affect	Democratic Quality	Delivery Quality
Log GDP per capita	1	0,5964	0,8390	0,3623	0,2679	0,6796	0,7710
Social support	0,5964	1	0,4619	0,4883	0,4315	0,5704	0,5064
Healthy life expectancy at birth	0,8390	0,4619	1	0,3090	0,1990	0,6068	0,6985
Freedom to make life choices	0,3623	0,4883	0,3090	1	0,6422	0,4847	0,5195
Positive affect	0,2679	0,4315	0,1990	0,6422	1	0,3308	0,3251
Democratic Quality	0,6796	0,5704	0,6068	0,4847	0,3308	1	0,8672
Delivery Quality	0,7710	0,506455	0,6985	0,5195	0,3251	0,8672	1

Table 4: Pearson Correlation

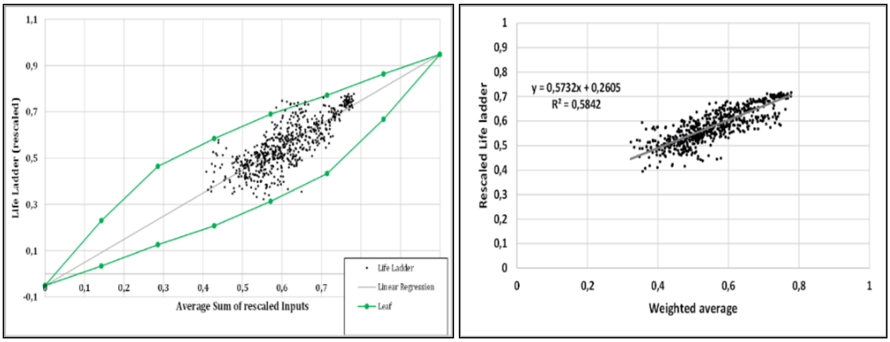

Figure 5: (left) Centred leaf; (right) S-weighted average

3.4 Comments

The results, summarized in the following table, exclude the possibility that the inputs selected are substitutes.

Hypothesis	Test result	Brief conclusion	Summary
Nonnegative intercepts of inputs	(see table 3)	The null hypothesis cannot be rejected for all the inputs but the GDP per capita	The inputs selected have an appropriate behaviour but the GDP per capita
Positive slopes of inputs	(see table 3)	The slopes are significantly positive for each input	Success
Each slope is lower than one	(see table 3)	All slopes are lower than one but GDP per capita and Health	Failure
Sum of slopes equal to one	(see table 3)	The sum is abundantly higher than one	Failure
Intercept Leaf=0	P-value = 0,0064	The null hypothesis can be rejected	Success only for an acceptable p-value of 0.1
Slope Leaf>0	P-Value <0,0001	The hypothesis is statistically significant	Success
Slope Leaf=1	P-value = 0,9589	The Hypothesis cannot be rejected [St.test F = 0,0026]	Success
Points in the Leaf	0%	The Leaf misses the 100% of the points	Failure
Points in the centred Leaf	98,57%	The Leaf centred misses the 1,43% of the points	The S-weighted average illusion fails to explain only the 1,43% of the results

Table 5: Summary

3.5 Further information and complementarity insights

The analysis presented here is the best result among many attempts that cannot be presented for space reasons. Other analysis included other variables in the data set quoted, the log of all the dependent variables, the log of the independent variables and their possible combinations. In particular, the test on the logarithm of all the variables tests implicitly the complementarity since the formula (6)

$$Out = \prod I_i^{a_i} + \varepsilon \quad (6)$$

Can be reconducted to (7)

$$\log(Out) = \sum a_i \log(I_i) \quad (7)$$

That can be treated with the model presented. Hence, the complementarity must be excluded as well.

4 Conclusions and implications

The substitutability of the elements impacting the happiness would entail both that the utility functions and the composite indicators should consider only the sum of these elements and not their reciprocal and relative proportions. When the perfect substitutability conditions are met: 1) mono-dimensional measurement are justified; 2) each indicator can be treated separately and, therefore, 3) policy maker can focus on different targets independently.

According the data analysed, there is neither substitutability nor complementarity. Hence, life ladder analysis cannot be satisfactorily performed by a utility function represented by a mere weighted sum of the inputs. This consideration can be extended to S-weighted-average-based indicators. It follows that a policy maker aiming to increase the life ladder, or a statistician aiming to measure it, must consider a higher complexity where the relationships among the inputs, and not their sum, whenever balanced, may explain happiness. In turn, this supports research on the idiosyncrasy and subjectivity of happiness promoting further research on regional or local levels.

References

[1] Stiglitz, J., Sen, A., & Fitoussi, J.-P. (2009). Report by the commission on the measurement of economic performance and social progress.

[2] United Nations. (2015). Transforming our world: the 2030 agenda for sustainable development . United Nations A/RES/70/1.

[3] Easterlin, R. (1974). Does economic growth improve the human lot? some empirical evidence. In R. David, & M. Reder, Nations and households in economic growth: essays in honor of Moses Abramovitz (pp. 89-125). New York: Academic Press.

[4] Helliwell, J., Layard, R., & Sachs, J. (2017). World happiness report 2017. New York: Sustainable development solutions network.
[5] Helliwell, J., Huang, H., & Wang, S. (2017). The social fundations of world happiness. In United Nations, World Happiness Report 2017 (pp. 8-47). United Nations SDSN.