Università di Pisa

Sant'Anna Scuola Universitaria Superiore Pisa

Consiglio Nazionale delle Ricerche

Università degli Studi di Cagliari

# Book of Short Papers SIS 2021

SIS2021 PISA, 21-25 GIUGNO 2021

Società Italiana di Statistica

Editors: Cira Perna, Nicola Salvati and Francesco Schirripa Spagnolo

PUBLISHED BY PEARSON

WWW.PEARSON.COM

*ISBN 9788891927361*

# Contents

# 6 Satellite events 1660

# An Application of Temporal Poset on Human Development Index Data

## *Un'applicazione del Temporal Poset ai dati dell'Indice di Sviluppo Umano*

Leonardo Salvatore Alaimo, Filomena Maggino and Emiliano Seri

**Abstract** Within the international debate on finding measures beyond GDP, the Human Development Index released by the United Nations Development Programme, has become a reference over the years. In order to get a synthetic view of human development, different aggregative procedures has been applied over time. The aggregative road to synthesis is however problematic, because it raises a number of conceptual and methodological issues. As a valuable alternative, in this paper we adopt a non-aggregative approach to synthesis over time, based on Partially Ordered Set Theory.

**Abstract** *Nel dibattito internazionale sulla ricerca di misure che vadano oltre il PIL, l'Indice di Sviluppo Umano sviluppato da* United Nations Development Programme*, è diventato un riferimento nel corso degli anni. Al fine di ottenere una visione sintetica dello sviluppo umano, sono state applicate nel tempo diverse procedure aggregative. La strada aggregativa verso la sintesi è tuttavia problematica, poichè pone una serie di problematiche concettuali e metodologiche. Come valida alternativa, in questo articolo adottiamo un approccio non aggregativo alla sintesi, basato sulla teoria degli insiemi parzialmente ordinati.*

**Key words:** Human Development Index, Synthesis of statistical indicators, Compensability, Partially Ordered Set - Poset

---

Leonardo Salvatore Alaimo
Corresponding author. Italian National Institute of Statistics - Istat, e-mail: `leonardo.alaimo@istat.it`

Filomena Maggino
Sapienza University of Rome e-mail: `filomena.maggino@uniroma1.it`

Emiliano Seri
Sapienza University of Rome e-mail: `emiliano.seri@uniroma1.it`

## 1 Introduction

Although it has been considered a *niche field*, the topic of synthesis of statistical indicators has a rich and varied scientific literature. The growing attention on this issue is, in a way, linked to the international debate on identifying measures that go *Beyond GDP*. From this perspective, the synthetic approach becomes the only one possible for a correct understanding of the phenomenon (7). The traditional statistical approach to synthesis is the so-called aggregative-compensative, according to which the synthetic measure is the result of the mathematical aggregation of the basic indicators (composite indicators). This has become the dominant framework over time. However, it poses a series of conceptual and methodological criticalities that have been highlighted in recent literature (3; 2). To try to overcome these limitations, research has focused on methods belonging to the so-called non-aggregative approach, in which the synthetic measure is obtained by not combining the basic indicators. Among those methods, the theory of partially ordered sets (poset theory) has become a reference over the years. This method, particularly suitable for the treatment of ordinal data, is useful even if we deal with indicators of different scaling levels. In this paper, we apply a poset-based synthesis method suitable for time series data (the so-called *temporal poset*) developed by (1) to the data of the synthetic measure elaborated by the United Nations Development Programme (UNDP), Human Development Index (HDI). This is a synthetic measure elaborated by the United Nations Development Programme (UNDP) and conceptually based on Sen's capabilities approach (8). HDI identifies three main dimensions, the basic capabilities crucial to human development: a long and healthy life, knowledge and a decent standard of living. The need to identify alternative aggregations (also UNDP posed the problem, changing the aggregation procedure, which was previously an arithmetic mean, in 2010 adopting a geometric one) is closely linked to a central issue in the context of composite indicators, the level of compensability or substitutability allowed between basic indicators. Generally, the basic indicators of a composite index are called *substitutable* if a deficit in one may be compensated by a surplus in another; on the contrary, the basic indicators are called *non-substitutable* if a compensation among them is not allowed. Consequently, an aggregation approach can be 'compensatory' or 'non-compensatory' depending on the adoption or not of compensation. The issue is not only methodological but also, and above all, conceptual. Let us consider HDI. If we admit full compensability, we implicitly affirm, for instance, that a surplus in the education dimension can compensate for a deficit in the economic one. This is, at least, highly questionable. On the other hand, if we affirm the non-compensability of the basic indicators, we risk flattening the results of our synthesis downwards (2). Finally, if we adopt a partially compensative method, i.e. allowing it "up to a certain point", the question would arise as to what is the permissible and tolerable threshold of compensability. The problem is not so much one the compensability as the aggregation, which generates a *flattening effect* (3) regardless of the method and, consequently, the level of compensability allowed. The method adopted in this paper not only addresses these issues but also offers a possible solution, as demonstrated by different empirical studies (1; 2; 3).

## 2 Data

As previously mentioned, the data used in this paper are those relating to the Human Development Index (http://hdr.undp.org/en/content/download-data). In particular, we refer to the time series from 2016 to 2019 of the four indicators that compose the three dimensions of the HDI: *life expectancy at birth*; *expected years of schooling*; *mean years of schooling* and *gross national income (GNI) per capita*. The indicators described are therefore those necessary to give the most complete and exhaustive representation of human development, understood as the advancement of human freedom, dignity and equality, and these include the broad comprehensive range of freedoms covering economic, social, political and civil areas. For a better illustration of the proposed methodology we only used the European countries, which correspond to our units in the four years considered.

## 3 Methodological aspects of the application

Poset supplies concepts and tools that appropriately adapt to the needs of synthesis. It is focused on *profiles*, which are the combinations of scores of each statistical units in the basic indicators considered, describing the status of units. This approach has some advantages and overcomes some limitations of the aggregative-compensative approach. We refer to the extensive literature on the basic aspects and definition of poset (4) and on the methodology about its use for data synthesis over time (1; 2). in this paper, we focus on describing the different steps leading to the construction of the synthetic measure according to this method.

First of all, the indicators must all have positive polarity[1]; where some have negative polarity, this must be reversed using a transformation. This is necessary to ensure that nodes in the highest positions of the Hasse diagrams will indicate better situations than those in the lowest positions. The system considered in our work is composed by 43 units (the European countries), 4 indicators and 4 temporal occasions (2016-2019). This is a three-way data array $\mathbf{Y} \equiv \{y_{ijt} : i = 1,\dots,43;\ j = 1,\dots,4;\ t = 1,\dots,4\}$ that can be seen as a set of 4 matrices of order $(43 \times 4)$, each of which represents a temporal slice of $\mathbf{Y}$. For each of the 4 matrices independently, we can calculate the incidence matrix[2] and construct the Hasse diagrams[3] , reported in Figure 1. Just the graphical representations give important information; it is evident, for instance, that the relationship structure of the system is different in the two times considered.

---

[1] Polarity is the sign of the relation between the indicator itself and the phenomenon

[2] The incidence matrix is a matrix $Z_P = (z_{ij}) \in \mathbb{Z}^{k \times k}$, where $|X| = k$ is the cardinality of $X$ and $z_{ij}$ is equal to 1 if $x_i \trianglelefteq x_j$, 0 otherwise, with $x_i, x_j \in X$. It defines the structure of comparabilities in the poset

[3] The *Hasse diagram* is the graphical representation of the directed acyclic graph representing the cover relation $\prec$: two elements are comparable $\trianglelefteq$ if a path connects them in the Hasse diagram.

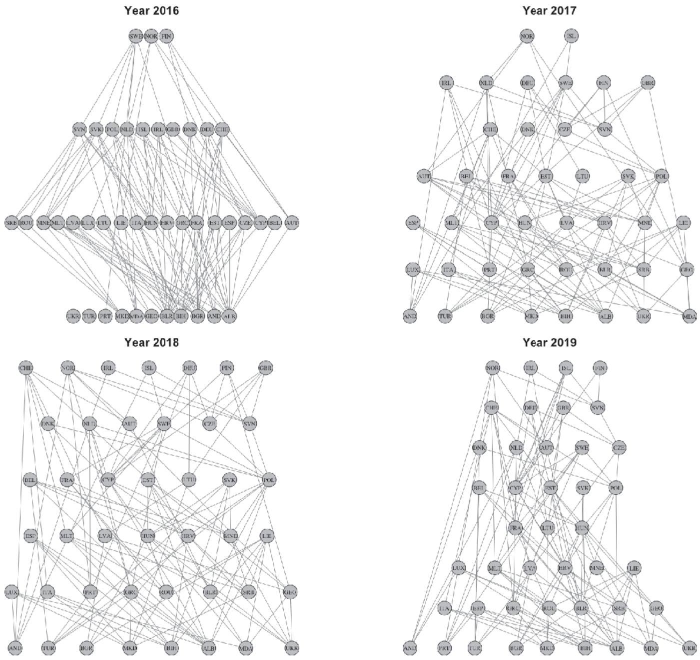


Fig. 1: Hasse diagrams of the European countries: years 2016 - 2019.

We want to obtain a synthetic measure. To do this, poset offer different possibility; in this paper, we use the so-called *average height*[4]. We can compute this measure for each of the 4 posets[5]. The results obtained allow an *intra-temporal* comparison of the units within the system. For example, we can affirm that Norway (NOR) is better than Italy (ITA) in 2017 (Figure 1). Anyway, it is impossible to make an *inter-temporal* comparison of units. For instance, Spain (ESP) worsens from 2018 to 2019, but we do not know why this happened. For instance, it may have happened that all the other countries had a very marked improvement in the indicators considered, while ESP could have increased slightly and been overtaken by the other units. Another possibility could be that Spain has been drastically reduced from 2018 to 2019 compared to the other units.

To make comparisons over time, we must *merge* the posets. Given two finite posets $\Lambda$ and $\Pi$, we merge them by setting $x \leqslant_{\Lambda\Pi} y$ if and only if one of the following conditions is valid (1):

1. $x, y \subset \Lambda$ and $x \leqslant_{\Lambda} y$;
2. $x, y \subset \Pi$ and $x \leqslant_{\Pi} y$;
3. $x \subset \Lambda$; $y \subset \Pi$ and $x \leqslant_{\Lambda\Pi} y$;

[4] For a definition and the methodological step of its calculation, please see (1; 2).
[5] We do not include the results in the paper for question of space. They can be required to authors.

4. $x \subset \Pi$; $y \subset \Lambda$ and $x \leqslant_{\Lambda\Pi} y$.

In other words, by merging the two posets we maintain their initial structures of comparability, adding other comparabilities that are an expression of the temporal comparison among the elements. In this way, it will be possible to make intertemporal comparisons. Moreover, in order to make it possible to compare posets with different sets of nodes and to anchor the the average height computation to a *common reference system*, we introduce an *embedded scale* (6), i.e. some benchmark profiles that form a scale of increasing levels embedded in the original poset. They are points that help anchoring the comparisons between profiles in the Hasse diagram and the average heights to a reference scale. We can merge the 4 posets in Figure 1 in one temporal poset and add a 5 levels embedded scale[6]. By calculating the average height of the resulted *temporal poset*, we can make comparisons over time of different units, using benchmarks as a common reference system for the different years. For instance, we can observe that ESP passes from a value of average height of 147 in 2018 to a value of 126 (see Table 1) and compare this trend with those of other countries.

## 4 Conclusions

HDI aims to allow policy makers evaluation of national policies. Thus, it is important to have a measure of the human development not affected by compensation in order to observe how the phenomenon behaves in all its dimensions. The proposed method tries achieving this goal, giving an easy-to-read representation of the HDI for each country in all the considered years. In the synthetic index proposed, we do not focus on the values of basic indicators and on their aggregation, but on the profiles of each country. This allows the overcoming of the compensability issue. The results, shown in Table 1, are free from the *flattening effect*, typical in the mean–based aggregation methods: countries with different profiles present different average heights.

## References

[1] Alaimo, L.S. Complexity of Social Phenomena: Measurements, Analysis, Representations and Synthesis. Unpublished doctoral dissertation, University

[6] The scale is defined as follows:

- *min*, with a profile given by the minimum value in all indicators;
- $B1$, with a profile given by the first quartile of all indicators;
- $B2$, with a profile given by the second quartile of all indicators;
- $B3$, with a profile given by the third quartile of all indicators;
- *max*, with a profile given by the maximum value in all indicators.

| European countries | 2016 | 2017 | 2018 | 2019 | European countries | 2016 | 2017 | 2018 | 2019 |
|---|---|---|---|---|---|---|---|---|---|
| **ALB** | 16.79 | 34.30 | 24.06 | 38.23 | **LIE** | 68.42 | 87.14 | 108.54 | 133.06 |
| **AND** | 38.46 | 21.45 | 45.66 | 73.93 | **LTU** | 89.73 | 113.87 | 125.06 | 136.52 |
| **AUT** | 103.32 | 115.61 | 127.81 | 112.37 | **LUX** | 81.81 | 132.79 | 124.02 | 147.55 |
| **BEL** | 91.96 | 112.81 | 135.16 | 153.46 | **LVA** | 69.08 | 77.92 | 89.58 | 101.79 |
| **BIH** | 11.89 | 11.30 | 22.34 | 34.39 | **MDA** | 7.37 | 15.00 | 26.59 | 19.37 |
| **BLR** | 40.58 | 48.48 | 56.35 | 49.04 | **MKD** | 4.83 | 9.37 | 14.58 | 20.21 |
| **BUL** | 20.67 | 21.09 | 15.19 | 32.80 | **MLT** | 67.49 | 82.11 | 95.86 | 111.37 |
| **CHE** | 148.50 | 152.79 | 162.71 | 169.67 | **MNE** | 41.31 | 48.21 | 55.83 | 62.58 |
| **CYP** | 58.87 | 72.00 | 85.14 | 92.05 | **NLD** | 107.88 | 120.87 | 133.65 | 150.29 |
| **CZE** | 105.42 | 100.15 | 109.45 | 120.12 | **NOR** | 141.52 | 152.60 | 161.00 | 168.86 |
| **DEU** | 148.20 | 155.77 | 162.63 | 169.65 | **POL** | 74.39 | 81.20 | 88.15 | 86.01 |
| **DNK** | 112.22 | 133.71 | 153.13 | 141.59 | **PRT** | 50.50 | 45.18 | 60.35 | 76.84 |
| **ESP** | 96.11 | 115.07 | 147.04 | 126.82 | **ROU** | 27.51 | 28.45 | 36.72 | 45.34 |
| **EST** | 110.41 | 105.13 | 131.72 | 117.48 | **SRB** | 13.07 | 23.44 | 32.55 | 33.13 |
| **FIN** | 139.34 | 149.19 | 159.02 | 168.48 | **SVK** | 89.20 | 113.96 | 141.33 | 158.64 |
| **FRA** | 83.44 | 96.38 | 109.34 | 123.21 | **SVN** | 85.42 | 99.49 | 113.16 | 127.56 |
| **GBR** | 132.78 | 145.19 | 157.91 | 168.23 | **SWE** | 121.37 | 133.05 | 141.32 | 167.13 |
| **GEO** | 28.65 | 41.81 | 54.89 | 78.56 | **TUR** | 9.79 | 18.93 | 29.05 | 39.74 |
| **GRC** | 63.24 | 76.59 | 90.50 | 104.13 | **UKR** | 7.66 | 17.81 | 18.64 | 29.92 |
| **HRV** | 52.09 | 58.33 | 64.10 | 70.01 | **MIN** | 1.00 | 1.00 | 1.00 | 1.00 |
| **HUN** | 54.33 | 55.62 | 66.78 | 74.20 | **B1** | 36.58 | 36.58 | 36.58 | 36.58 |
| **IRL** | 134.53 | 149.17 | 163.48 | 165.04 | **B2** | 73.65 | 73.65 | 73.65 | 73.65 |
| **ISL** | 143.10 | 153.78 | 157.02 | 160.45 | **B3** | 128.43 | 128.43 | 128.43 | 128.43 |
| **ITA** | 60.98 | 77.92 | 98.77 | 121.06 | **MAX** | 177.00 | 177.00 | 177.00 | 177.00 |

Table 1: Average height values: European countries; years 2016-2019.


of Rome ”La Sapienza”, Rome, Italy (2020).

[2] Alaimo, L.S., Arcagni, A., Fattore, M ., Maggino, F. Synthesis of Multi-indicator System. Soc. Ind. Res. (2020) doi: 10.1007/s11205-020-02398-5.

[3] Alaimo, L. S., Maggino, F. Sustainable development goals indicators at territorial level: Conceptual and methodological issues—The Italian perspective. Soc. Ind. Res., 147(2), 383–419. (2020) doi: 10.1007/s11205-019-02162-4.

[4] Bruggemann, R., Patil, G.P. Ranking and prioritization for multiindicator systems: Introduction to partial order applications. Springer, Dordrecht. (2011).

[5] D'Urso, P., Alaimo, L. S., De Giovanni, L., Massari, R. Well-being in the Italian regions over time. Soc. Ind. Res. (2020) doi: 10.1007/s11205-020-02384-x.

[6] Fattore, M. Non-aggregated Indicators of Environmental Sustainability. Silesian Statistical Review/Slaski Przeglad Statystyczny, 16(22), 7–22. (2018).

[7] Maggino, F., Alaimo, L. S. Complexity and wellbeing: measurement and analysis. In: Bruni, L., Smerilli, A., De Rosa D. (eds.) A Modern Guide to the Economics of Happiness, pp. 113–128. Edward Elgar Publishing, Northampton (2021).

[8] Sen, A. Commodities and Capabilities. Oxford University Press, Oxford. (1999).