

A composite indicator for the measurement of infrastructures endowment - Experiences in Istat

Claudia Brunini, Francesca Paradisi

Abstract The present work focuses on aspects of methodological innovation introduced during the activities conducted by the Istat Research Unit related to the computation of a mono-dimensional measure. The innovations are about the techniques for the synthesis of basic indicators and the stages that precede the application of the method. The work also shows the results obtained by the application of an original technique of synthesis on a set of urban transport infrastructures indicators.

Keywords: Composite Indicators, Environmental Statistics

1 Foreword

The authors, as members of a research unit of Istat, operating under the project "Territorial Information for the Structural Policy 2001-2008", are dealing with issues related to the infrastructures of a territory since 2002.

"Le Infrastrutture in Italia" (Istat, 2006), represents a crucial first goal on the activity of the building of a system of indicators and practicality of infrastructures. The system, built on respect of fundamental characteristics (official sources; indicators divided according to the classification adopted; nuts3 as territorial desegregation; time series available since 1996; temporal and spatial homogeneity), offers a very wide overview of data and indicators. However the system of indicators, while providing rich and detailed information, can make the reading and analyzing too structured. Because of these facts, it has been considered the possibility of using a mono-dimensional measure that

¹

Claudia Brunini, Istat; brunini@istat.it:

Francesca Paradisi, Istat; paradisi@istat.it:

summarizes the information to make it immediately visible and interpretable. During the years, the studies on the argument have facilitated the creation of numerous papers. The first one, which was presented at the Sixth National Conference of Statistics (Brunini, Messina, Paradisi, 2002), and others (Brunini, Messina, Paradisi, 2003) have as main goal the analysis of existing literature. They are also intended as a critical reading of the methods for the synthesis of indicators, in particular, the many standardization and synthesis methods² have been the subject of extensive studies, which have highlighted the strengths and weaknesses and the effects on the output.

In the following contributions the Research Unit has proposed alternative and innovative ways to build a composite indicator. The Istat Seminar *Informazione statistica, territorio e attività di policy* (Brunini, Messina, Milazzo, Paradisi, 2004), was an opportunity to present for the first time an alternative synthesis method. The method was also illustrated at the MIUR seminar *Metodi e strumenti per la produzione di informazioni statistiche territoriali*. Istat published the transport infrastructures endowment index and the facilities intermediation endowment index, both obtained applying the method (Istat, Rapporto Annuale 2006). The finished version of the method was published in *Rivista di Statistica Ufficiale* (Brunini, Paradisi, 2006) and in *Indicatori e modelli statistici per la valutazione degli squilibri territoriali* (Brunini, Paradisi, 2009).

Between 2007 and 2008 the research activity continued with the contribution of the Istat Methodological Unit. The partnership, formalized in a working group³, has been very fruitful and it has contributed to the development of a new synthesis technique, which relies on many of the principles already formulated. Principles include the non-substitutability of the indicators, as unique referents of the concept, and the horizontal variability that triggers the idea of penalization to be applied to the simple arithmetic mean of the indices. The group has produced a volume that contains a summary measure of the endowment of health services in Italian regions (Istat, 2009). It is an important arrival point, since it represents one of the few publications in which the Istat decided to disclose a measure of synthesis. This milestone had already been touched by the Research Unit in 2006, when it was published in the Annual Report the infrastructures endowment index for specific sector areas.

The present work focuses on aspects of methodological innovation introduced during the activities conducted by the Research Unit. The innovation is about the techniques for the synthesis of basic indicators and the stages that precede the application of the method. The paper also shows how the method elaborated by the authors works on a set of urban transport infrastructures indicators.

2 Elements of Innovation

² Synthetic methods were classified into four groups, according to methodological approach: the first one includes methods that produce ordinal measures, the second one makes use of multivariate analysis techniques, the third one is based on distances or quadratic functions and, in the end, the fourth one use as aggregate function, the arithmetic mean or other linear function.

³ The working group (created with resolution Istat n. 180/07/DCET 8 May 2007), was coordinated by Claudia Brunini and composed by Alessandra Burgio, Alessandro Solipaca, Matteo Mazziotto, Adriano Pareto, Antonino Messina, Sara Basso, Emanuela Fidale and Rosario Milazzo.

The Research Unit has achieved important and innovative goals during his activity.

The first innovative element introduced is about the definition of infrastructure, it is articulated in order to expand the traditional vision representing the endowment only in a physical sense. The identification of the set of indicators is designed so that everyone is non-replaceable and representative of the extent of each sub-group the multi-dimensional concept being measured is divided. The indicators can be standardized in order to maintain or cancel the effect of implicit weighting due to differences in variability (vertical variability). Even the concept of horizontal variability, seen as variability in the indicators in the same territorial unit, represents an innovative element in the reflection group on the methods. Under the assumption of no substitution of variables, a high degree of dispersion between the standardized indicators means a great imbalance in the availability of resources. The fact, even negative, led to introduce in the method of synthesis a penalization factor as direct function of the horizontal variability. This factor adjusts the arithmetic mean of the standardized indicators.

2.1 *A broader definition for infrastructures endowment*

Traditionally the level of infrastructures endowment of an area is expressed in physical terms and refers to the presence of only physical resources. In the area of health infrastructure, for example, is defined as supply of hospitals and beds, or in the education area, as supply of school buildings and classrooms.

The introduction of a broader concept of infrastructure endowment represents a first and crucial factor of innovation in the analysis. Besides the physical resources in the narrow sense are considered also instrumental means and human resources. The health area is then integrated with indicators of the presence of medical personnel and diagnostic equipment, while education is enriched with measures concerning the presence of teaching staff and various facilities (canteen, school, etc.).

2.2 *Indicators non-replaceable and representative*

The selection of indicators is a process of considerable importance because it is one that most affects the value and quality of composite index. The choice is to give priority to the conceptual criteria although taking full account of the statistics relations existing between the variables.

The set of indicators is highly dependent on the availability of sources and their quality. After a careful survey of existing data from official⁴ sources and uniformity in time and space, we proceed with the selection of variables, chosen so that each dimension of the concept is described as best as possible. In order to accomplish this, we employ well-

⁴ The official sources allow to minimize the non-response problem and ensure a good level of homogeneity between the surveys.

known experts. The result is a set of representative indicators of the phenomenon, among them non-replaceable, because able to measure a specific dimension of the concept represented.

The effective congruence with the theoretical dimensions which goes on to be described is verified by careful conceptual reflection that takes full account the informative capacity of the indicators, as well as by statistical techniques that helps to determine the relations between them.

2.3 *Implicit weight due to vertical variability*

As known the aggregative procedures for synthesis assume that the elementary indicators are summarized and therefore they should be expressed in the same unit. This is achieved through the process of standardization. If the objective of standardization is only to obtain pure numbers it is desirable that it should affect as little as possible in the characteristics of the distribution. However, extensive experiments have shown that in literature there is no standardization criteria able to leave unchanged the distribution of the index and that the goal for which to strive, rather, is to maintain as much control as possible of the effects caused by this transformation. Here we consider two criteria of standardization⁵ emblematic for their differences, as the first, dividing the values by a constant, preserves the variability of the spatial distribution and relative distances between each unit, the second to the contrary, reporting the distribution of the variable to a normal Gaussian with mean and variance constant, affects both these elements.

The first of the two methods maintains the relative distance between the units. It also assigns a weight proportional to its variability (implicit weight). The choice of this method of standardization has both a statistical and a conceptual justification. If the indicators were already in the same unit of measure, the researcher does not operate on them with any further processing before aggregating them and proceed to synthesize all the information without concerning about the original distribution. Moreover, from a conceptual point of view, if the objective of the analysis is to measure the endowment of an area in comparison with all other territorial units, then why change the original distances between units and bring all indicators in the same normal distribution, effectively nullifying the extra information on the disparity between the units?

On the other hand, the second method of standardization, produces a series of indicators with mean zero and variance equal to one, actually nullifying the effect due to spatial variability of the series, giving the same weight to each indicator in the synthesis, and changing substantially the original spatial distribution.

⁵ The array of standardized indicators $X = (x_{ij})$ is of n rows (territorial units) and m columns (indicators). The method 1 is the division of each observation x_{ij} for the target value of the j -th indicator; the target values vector is obtained from a cluster analysis performed on the matrix X ; the centroid of the cluster with better performance is the value used to standardize. The method 2 is the division between the deviation of each observation x_{ij} and the arithmetic mean of the series for the mean square deviation of the territorial series.

2.4 A factor penalty proportional to horizontal variability

All the alternative synthesis technique proposed by the Research Unit use for aggregating indicators the arithmetic mean, that is the most used criterion in literature for building a composite infrastructures endowment index. The innovative element is represented by the introduction of a penalty factor to correct the arithmetic mean. The factor is a function of variability between the indicators (so-called horizontal variability). The idea behind the correction is that, under the assumption that all indicators are an essential aspect of the concept, none of them may be lacking or replaced by a surplus of another.

So the Imbalances Difference Weighted Method corrects the mean with a direct function of the horizontal variability⁶. Moreover the method wants consider the local context into which the unit is placed: for this reason the penalty factor is multiplied by a weight⁷ that reduces his action in case the high variance among the indicators is widespread across regions and therefore be considered almost a 'physiological elements' of the context.

3 Application to urban transport infrastructures

The Imbalances Difference Weighted Method (IDW method)⁸ was applied to a set of 9 indicators on urban transport infrastructures (Istat, 25 marzo 2009) in the municipality⁹.

⁶ The function penalizes the territorial units that, compared to the others, have a greater disparity between the standardized indicators (high horizontal variability), as expression of imbalance in the allocation of resources.

⁷ The weights work on each element of the simple difference average in function of the vertical variability. The greater the difference between the regions is, the lower the weight assigned to the difference between the pair of indicators themselves. Thus, units with dissimilar indicators, in presence of moderate spatial variability, breaking the "law" governing the environment, are penalized more than they do in presence of high spatial variability.

⁸ The Imbalances Difference Weighted Method, with respect to the i-th territorial unit, took the following formula:

$$M_{v_i}^c = M_{v_i} - \left(M_{v_i} \cdot \Delta_i^{rel} \right) \quad (1)$$

where: $M_{v_i} = \frac{\sum_{j=1}^m v_{ij}}{m}$

is the arithmetic mean of the standardized indicators;

The indicators are:

density of railways, tramways, trolleys , subway networks, cable cars network (km per 100 km² of municipal surface);

density of bus networks (km per 100 km² of municipal surface);

availability of buses, trams, trolleys, subway cars (number per 10.000 inhabitants);

density of bus stops, tram, trolleys, subway stations (number per km² of municipal surface);

density of limited traffic areas (ZTL) (km² per km² of municipal surface);

$$\Delta_i^{rel} = \frac{\sum_{s=1}^m \sum_{t=1}^m |v_{is} - v_{it}| p_{st}}{\sum_{s=1}^m \sum_{t=1}^m p_{st}} \cdot \frac{1}{m M_{v_i.}} \quad \text{is the relative average simple difference}$$

weighted.

The weight p_{st} is 1 minus the arithmetic mean, of n regions, of the relative average simple difference for two generic indicators:

$$p_{st} = 1 - \left(\sum_{i=1}^n \Delta_i^{rel} \cdot \frac{1}{n} \right) = 1 - \left(\sum_{i=1}^n \frac{|v_{is} - v_{it}|}{(v_{is} + v_{it})} \cdot \frac{1}{n} \right)$$

The expression (1), after simple algebraic passages, becomes:

$$M_{v_i.}^c = \frac{\sum_{j=1}^m v_{ij}}{m} - \left(\frac{\sum_{s=1}^m \sum_{t=1}^m |v_{is} - v_{it}| p_{st}}{m \cdot \sum_{s=1}^m \sum_{t=1}^m p_{st}} \right) = M_{v_i.} - \frac{\Delta_i}{m} \quad (2)$$

$$\text{where, } \Delta_i = \frac{\sum_{s=1}^m \sum_{t=1}^m |v_{is} - v_{it}| p_{st}}{\sum_{s=1}^m \sum_{t=1}^m p_{st}}$$

$$\text{with: } 0 \leq M_{v_i.}^c \leq M_{v_i.}$$

availability of pedestrian areas (m² per 100 inhabitants);
parking pay for 1.000 cars;
parking exchange for 1.000 cars;
density of cycle track (km per km² of municipal surface).

Every indicator concerns the allocation of a specific kind of urban transport infrastructures, because of this they are non-replaceable. The indices are poorly correlated ($R < 0.35$) except for the first four, that are all related to public urban transport. Among these the highest correlation is 0,69, detected between the density of bus networks and the availability of buses, trams, etc.

The Table 1 shows the value of the Urban Transport Infrastructures Endowment Index (Italy = 100) for each of 111 municipality in 2007. The index is obtained with the application of the Imbalances Difference Weighted Method. The table shows also the rank occupied by every territorial unit according to the Index value.

The ranking obtained with the Imbalances Difference Weighted Method is not significantly different from that one obtained with the arithmetic mean. On 111 municipality, 39 occupy the same placing in both rankings, 45 have moved of only one rank and only three municipalities (Treviso, Pisa, Como) of 6 or more ranks. According to the frequency distribution of the two class index ¹⁰, there are only 7 municipalities that fall into a different class.

The map shows that 31 municipalities have an index that is less than half the national average (Italy =100), 29 fall into the lower middle class (50-100), 18 recorded a value between 100 and 150, and 33 have a value more than one and half times the Italian average; among these there are some municipality with insomuch value as to exceed more than three times the Italian average (Pavia, Bergamo, Firenze and Torino) or more than two times as Bologna, Biella, Brescia, La Spezia, Padova, Milano, Vercelli and Mantova. The municipalities with the lower index values (less than a quarter of national average) are Ragusa, Tempio Pausania, Olbia, Tortoli, Enna, Lanusei, Sanluri, Villacidro, Carbonia, Iglesias. Apart from few exceptions, the map shows the traditional great difference between the northern regions, with high and medium-high urban transport infrastructures endowment, and Central and Southern areas, where are observed a significantly lower index values.

Table 1: Urban Transport Infrastructures Endowment Index - Year 2007

Municipality	IDW method		Arithmetic mean	
	Index (Italy=100)	Rank	Index (Italy=100)	Rank
Pavia	397,94	1	445,62	1
Bergamo	384,47	2	425,76	2
Firenze	324,34	3	326,78	4
Torino	324,21	4	328,84	3
Bologna	271,73	5	274,45	6
Biella	254,19	6	291,20	5
Brescia	246,37	7	258,09	7

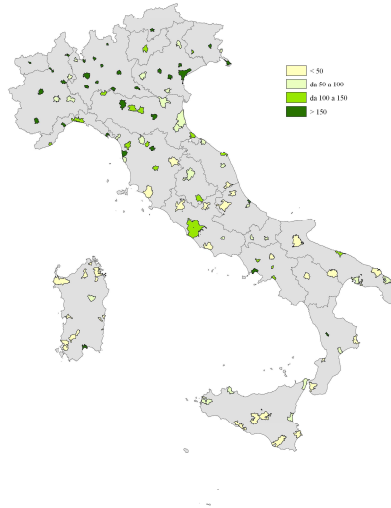
¹⁰ The classes are four, symmetric to the value Italy = 100: low (<50), medium low (50-100), medium high (100-150) and high (> 150).

La Spezia	237,76	8	246,73	8
Padova	228,55	9	242,87	10
Milano	223,01	10	243,66	9
Vercelli	218,90	11	242,55	11
Mantova	209,47	12	217,09	12
Napoli	196,96	13	209,13	14
Cosenza	191,20	14	209,64	13
Aosta	190,59	15	194,51	16

⋮

Latina	29,83	97	31,56	97
Siracusa	28,97	98	30,78	99
Matera	27,78	99	30,83	98
Sassari	26,96	100	30,77	100
Caltanissetta	26,13	101	29,17	101
Ragusa	24,22	102	26,12	102
Tempio Pausania	22,76	103	24,88	105
Olbia	22,53	104	25,10	104
Tortolì	21,89	105	25,19	103
Enna	21,03	106	24,43	106
Lanusei	20,53	107	22,61	108
Sanluri	18,50	108	23,21	107
Villacidro	17,79	109	19,48	109
Carbonia	16,68	110	18,65	110
Iglesias	5,00	111	5,71	111

Figure1: Urban Transport Infrastructures Endowment Index (Italy=100) – Year 2007



References

1. Brunini C., Messina A., Paradisi F. (2002). *Le infrastrutture delle province italiane: metodi e sperimentazioni*. VI Conferenza Nazionale di Statistica, Roma.
2. Brunini C., Messina A., Paradisi F. (2003). *Banca dati delle infrastrutture: un utile strumento per la valutazione*. AIV Congress (*Il contributo della valutazione alle politiche di sviluppo locale*), Reggio Calabria.
3. Brunini C., Messina A., Milazzo R., Paradisi F. (2004). *Infrastrutture e territorio: una proposta di metodo di sintesi di indicatori*. Istat seminar (*Informazione statistica, territorio e attività di policy*), Roma, 2th February 2004. MIUR seminar (*Metodi e strumenti per la produzione di informazioni statistiche territoriali*), Roma, 'Roma Tre' University, March 2004.
4. Brunini C., Paradisi F. (2006). *Una tecnica alternativa per la determinazione di una misura sintetica di infrastrutturazione*. Rivista di Statistica Ufficiale. Istat, n. 2/2006.
5. Brunini C., Paradisi F. (2009). *Costruzione di un indice sintetico della dotazione infrastrutturale delle province italiane: problemi e nuovi approcci*. In *Indicatori e modelli statistici per la valutazione degli squilibri territoriali*, Franco Angeli 2009.
6. Istat (2006). *Le infrastrutture in Italia*. Istat, Informazioni. n. 7 -2006.
7. Istat (Rapporto Annuale 2006). *Rapporto Annuale. La situazione del Paese nel 2004*. Istat, Roma.
8. Istat (25 marzo 2009). *Indicatori dei trasporti urbani. Anno 2007*. Approfondimenti del 25 marzo 2009.

9. Istat (2009). La dotazione di infrastrutture e servizi nella sanità. Istat, Informazioni n. 9 - 2009.