

A New Perspective in Comparative Analysis of Information Society Indicators

Pavle Sicerl

Law School, University of Ljubljana and SICENTER, Brajnikova 19, 1000 Ljubljana, Slovenia

Tel: +386 61 1501510; fax: +386 61 1501514

Pavle.Sicerl@sicenter.si

Keywords: time distance, S-distance, two-dimensional comparison in time and indicator space

Edited by: Cene Bavec and Matjaz Gams

Received: October 8, 1999

Revised: November 23, 1999

Accepted: December 12, 1999

The analysis of information society indicators can be enriched by supplying a new view of data that can provide new insight from existing data. The slowdown of growth of Internet hosts per 10000 inhabitants in Slovenia after mid-1997 increased the time lag of Slovenia behind leading Finland from 3 years at the end of 1996 to nearly 5 years by August 1999. Time distance methodology is used as a presentation and communication tool to raise awareness of the problem and its consequences in simple understandable terms and to signal the need for an in-depth analysis and action.

1 Introduction

Problem:

In Slovenia, after a very high rate of growth in the indicator Internet hosts per 10000 inhabitants until mid-1997, such growth slowed down substantially. One can describe the facts in various ways and with various measures.

Objective:

To make the government, other agents and general public aware of these developments and signal the need for immediate action to correct them.

Method:

Time distance will be used as a presentation and communication tool to raise awareness of the problem and its consequences in simple widely understandable terms. Since this method can be a useful addition to existing methods of analysing differences between compared units in many fields, a further illustration is provided for the case when the benchmark for comparison is the average value of the analysed indicator for EU15.

statistical measure for operational use, amendments to the present state-of-the-art are needed on two levels: conceptual and analytical.

First, a broader theoretical framework is required. The conventional approach does not realise that, in addition to the disparity (difference, distance) in the indicator space at a given point in time, in principle there exist a theoretically equally universal disparity (difference, distance) in time when a certain level of the indicator is attained by the two compared units. Second, a statistical measure S-distance has been defined to suggest a possibility how the broader concept and reference framework can be measured in operational terms. The aim is to provide new insights from existing data due to an added dimension of analysis and thus to complement conventional statistical measures.

Time distance in general means the difference in time when two events occurred. We define a special category of time distance, which is related to the level of the analysed indicator. The suggested statistical measure S-distance measures the distance (proximity) in time between the points in time when the two compared series reach a specified level of the indicator X. The observed distance in time (the number of years, quarters, months, days, minutes, etc.) is used as a dynamic (temporal) measure of disparity between the two series in the same way as the observed difference (absolute or relative) at a given point in time is used as a static measure of disparity [1,2,3].

2 Methodology: time distance concept and statistical measure S-distance

The time perspective, which no doubt exists in human perception when comparing different situations, is systematically introduced both as a concept and as a quantifiable measure. Since events are dated in time, in time series comparisons, regressions, models, forecasting and monitoring, the notion of time distance always existed as a "hidden" dimension. In order to systematise and formalise the approach and define an appropriate

For a given level of X_L , $X_L = X_i(t_i) = X_j(t_j)$, and the S-distance, the time separating unit (i) and unit (j) for the level X_L , will be written as

$$S_{ij}(X_L) = \Delta T(X_L) = t_i(X_L) - t_j(X_L)$$

where T is determined by X_L . In special cases T can be a function of the level of the indicator X_L , while in general it can be expected to take more values when the same level is attained at more points in time, i.e. it is a vector which can in addition to the level X_L be related to time. Three subscripts are needed to indicate the specific value of S-distance: (1 and 2) between which two units is the time distance measured and (3) for which level of the indicator (in the same way as the time subscript is used to identify the static measures). In the general case also the fourth subscript would be necessary to indicate to which point in time it is related (T_1, T_2, \dots, T_n).

The sign of the time distance comparing two units is important to distinguish whether it is a time lead (-) or time lag (+) (in a statistical sense and not as a functional relationship):

$$S_{ij}(X_L) = -S_{ji}(X_L)$$

Using the comparison between two units it can be shown that the generic concept of time distance goes together very naturally with the existing concepts of static disparity at a given point in time and the notion of the growth rate over time. Table 1 provides a schematic example for such comparisons for a given indicator. Row one is the most frequently used type of comparative analysis; levels of the indicator at a given point in time are compared. In such comparison two points are used, for each of them we have three elements of information: (i) the respective level of the indicator, (ii) to which unit it belongs, and (iii) at what time it happened. In this case unit as well as time (since it is constant for static comparison) serve as identifiers, while the levels are used to calculate the static difference. Row two compares two levels of the indicator for each unit at two points in time, separately for each unit, which means that one calculation indicated in row two refers to unit 1, and another to unit 2. The simplest example would be growth rate for unit 1 and growth rate for unit 2. Here the unit is the identifier, while the numerical values on levels and time are used in calculating this measure.

These two steps are standard procedures. The first one represents the static type of comparison; the second one measures the dynamic properties of the indicator for each unit separately. Following the same logic, for the novel statistical measure S-distance in row three level is the same, level and unit serve as identifiers, and time is used for calculating time distance. It is remarkable that the notion of time distance, which can be in principle developed from the same information used in steps one and two, has not been developed theoretically and as a standard statistical measure.

	TIME	UNIT	LEVEL	Measure
TIME	same	2	2	static difference
UNIT	2	same	2	change over time
LEVEL	2	2	same	time distance

Table 1. Points of comparison for static difference, change over time and time distance (two units)

While there may be different problems involved in the calculation of these three types of measures, in terms of availability and comparability of data, in principle these three types of measure can be integrated into a formally consistent analytical framework. There are alternative ways of doing this, following from the distinction between backward looking (ex post) and forward looking (ex ante) time distances. They relate to different periods, past and future, the first belongs to the domain of statistical measures based on known facts, the second is important for describing the time distance outcomes of the results of alternative policy scenarios for the future. Looking backwards, ex post or historical time distance indicates how many years ago the more developed unit experienced a specified level of the indicator of the less developed unit at a given point in time [3]. A very important relationship shows that, ceteris paribus, time distance is a decreasing function of the magnitude of the growth rate of the indicator. This conclusion shows that the S-distance as a dynamic (temporal) measure of disparity offers a perspective which may be quite distinct from that provided by static measures.

This new view of the information is using level(s) of the variables as identifiers and time as a focus of comparison and numeraire. This approach and the broad range of its possible applications is much more complex and general, but the time distance is the priority choice because of its intuitive nature, and the importance of the time dimension in semantics of describing various situations in real life and forming our perceptions about them. In this paper only the application to comparison of one indicator between several units will be used. However, the approach has been generalised to complement conventional measures in time series comparisons, regressions, models, forecasting and monitoring, and to analysis of single time series [3] and to variables other than time [4]. In all such applications it can provide from existing data new insights due to an added dimension of analysis.

3 Data and results for Slovenia, EU15 countries and candidate countries

Data on Internet hosts per 10000 inhabitants used relate to the period end of 1993-August 1999 [5,6,7]. At present is the measurement and empirical analysis of information society indicators beset with problems. It is stated that the single most important obstacle to effective data collection is the lack of standardised definitions of information technology and the exclusion of important

costs associated with its use, like personnel and training expenses. A further weakness is the relative absence of systematic information how information technology is actually being used [8]. In addition to these general obstacles there may be also some specific reasons that the slowdown of the increase in Internet hosts per capita in Slovenia in the last two years shown in RIPE data may have been exaggerated [9]. We shall proceed by analysing the available RIPE data, yet there should be an

appropriate caution about possible inaccuracy in the available data. Comparative analysis of the differences among countries can be presented in two dimensions. The conventional static differences at a given point in time are in this paper complemented by the time distance dimension. Time distance in Table 3 is for practical reasons calculated for the levels of the indicator for those countries, which are behind Slovenia, and for the level of Slovenia for the countries, which are ahead of Slovenia.

	1993	1994	1995	1996	1997	1998	Aug. 1999
LUX	7.4	12.5	46.0	85.2	113.4	182.3	218.8
DAN	16.1	35.4	96.9	203.3	321.1	571.1	608.3
BEL	7.0	17.3	30.2	64.0	104.8	202.4	307.5
AUT	18.9	34.0	66.3	110.2	134.4	214.0	235.7
DEU	13.7	24.4	58.0	84.4	137.7	177.0	186.3
FRA	9.3	14.4	26.0	40.6	60.7	84.2	106.4
NED	28.6	55.8	110.8	173.4	249.2	395.1	481.9
ITA	2.9	5.0	13.1	25.8	44.2	64.5	96.4
SVE	47.0	84.8	164.1	269.0	394.0	429.9	569.5
UK	19.1	38.7	75.1	122.4	167.3	247.4	272.2
FIN	65.2	133.9	416.7	612.1	945.8	902.6	930.5
IRL	6.5	15.3	37.3	74.2	109.3	155.6	181.7
ESP	3.6	7.0	13.1	28.8	49.9	78.2	94.2
PRT	3.6	5.1	11.9	23.6	42.7	56.3	67.4
GRE	1.7	3.4	7.4	16.0	26.7	47.1	63.5
SLO	3.1	8.2	28.3	69.5	98.2	115.3	116.3
CZE	4.3	10.1	21.1	39.6	55.2	83.6	101.8
SVK	0.7	2.6	5.6	14.8	27.0	41.0	48.3
HUN	3.0	6.6	15.4	29.2	66.7	87.8	106.2
POL	1.3	2.8	6.0	13.7	22.9	32.4	42.9
EST	2.9	7.7	24.1	54.3	108.4	151.2	180.8
ROM	0.0	0.2	0.8	3.5	6.0	9.9	14.1
LIT		0.3	1.2	4.7	10.9	26.0	32.7
LAT	0.2	2.0	5.2	23.1	28.6	54.4	63.7
BG	0.0	0.2	1.3	4.0	8.2	12.2	18.3
EU15	12.2	23.6	50.5	78.6	124.3	171.1	199.3

Table 2: Data on Internet host density per 10000 inhabitants

Source: International Telecommunication Union Database, Geneva 1998 for 1993-1997 [5]; RIPE [6] in RIS [7] for 1998 and 1999.

In Tables 2 and 3 the countries are sorted by the level of GDP per capita (at purchasing power standards) in 1997. Obviously, the Internet hosts per capita are not firmly correlated with GDP per capita. In 1996 Slovenia was occupying a comfortable comparative position in terms of Internet hosts per capita: it was lagging less than 3 years behind Finland as the leading country, and was ahead of several EU countries, i.e. Belgium, France, Italy, Spain, Portugal and Greece. The last four mentioned countries had substantially lower values than Slovenia.

The slowdown of growth rate in this indicator for Slovenia after mid-1997 led to a quick deterioration of the comparative situation of Slovenia. By August 1999

the lag behind Finland increased to nearly 5 years. Namely, in case of indicators with high rates of growth the situation can change very quickly, as distinct from the fields where the rate of change is slow. Figure 1 provides visualization of these changes. Tables 2 and 3, and Figure 1 compare Slovenia with EU15 countries and the nine candidate countries from Central and Eastern Europe.

One could also speculate what would be the situation if the rate of growth for the period 1997-August 1999 would continue until the end of 2000 (this should not be interpreted as projections).

	1994	1995	1996	1997	1998	Aug. 1999
LUX	-0.9	-0.5	-0.4	-0.5	-1.0	-1.5
DAN	#N/A	-1.4	-1.5	-2.0	-2.8	-3.4
BEL	-0.9	-0.2	0.1	-0.2	-0.9	-1.5
AUT	#N/A	-1.4	-0.9	-1.3	-1.8	-2.3
DEU	#N/A	-0.9	-0.6	-0.7	-1.4	-2.0
FRA	#N/A	0.1	0.7	1.2	1.5	1.1
NED	#N/A	#N/A	-1.8	-2.2	-2.9	-3.5
ITA	0.6	0.8	1.1	1.6	2.1	1.7
SVE	#N/A	#N/A	-2.4	-2.8	-3.6	-4.2
UK	#N/A	-1.5	-1.2	-1.5	-2.2	-2.7
FIN	#N/A	#N/A	-2.9	-3.5	-4.3	-4.8
IRL	-0.8	-0.4	-0.1	-0.3	-0.9	-1.4
ESP	0.2	0.8	1.0	1.5	1.7	1.7
PRT	0.6	0.8	1.2	1.7	2.3	2.6
GRE	0.9	1.2	1.6	2.1	2.5	2.7
SLO	0.0	0.0	0.0	0.0	0.0	0.0
CZE	-0.3	0.4	0.7	1.4	1.5	1.4
SVK	#N/A	1.5	1.7	2.1	2.7	3.1
HUN	0.3	0.6	1.0	1.1	1.4	1.1
POL	#N/A	1.4	1.7	2.3	2.9	3.2
EST	0.1	0.2	0.4	-0.2	-0.8	-1.4
ROM	#N/A	#N/A	2.9	3.4	3.9	4.3
LIT	#N/A	#N/A	2.7	2.9	3.1	3.5
LAT	#N/A	1.6	1.3	2.0	2.4	2.7
BG	#N/A	#N/A	2.8	3.0	3.8	4.1
EU15	#N/A	0.8	0.3	0.6	1.2	1.8

Table 3: Time distance between compared countries and Slovenia, S-distance in years: - time lead, + time lag, Slovenia=0
Source: Own calculation based on data in Table 1.

If no action would be taken and such slowdown would continue until the end of 2000, a further deterioration of the relative position of Slovenia for this indicator would take place. Slovenia would within a period of only a few years move from a comfortable position near the EU15 average in 1996 (despite being more than 30 per cent below the average EU15 level of GDP per capita) to a position where the lag behind the forerunner Finland would be already 6 years. The lag behind Sweden, Denmark and Netherlands would be around 5 years, France, Italy, Spain and Greece would surpass or catch up with Slovenia, and only Portugal out of the EU15 countries would be still behind it.

Time distance seems to be an excellent way of presenting the danger of a rapidly deteriorating situation, which everybody can understand, and to signal that an in-depth analysis and corresponding actions are necessary. Some other conventional measures may not provide such warning. E. g., static comparison showed that in 1996 Finland had 8.8 times the number of Internet host per capita in Slovenia, and in 2000 it would be 6.6 times. Time distance adds a qualitatively different conclusion.

Similar consequences can be seen from comparison with selected Central and Eastern European countries. In 1996 Slovenia was with Estonia a clear leader in the region for the indicator Internet hosts per capita. In the meantime Estonia moved ahead, and the gap would widen if the present trends would continue. By August 1999 Slovenia is lagging behind Estonia for more than 1 year.

The quality of time distance measure, being transparent and easy to perceive and understand, can be even more appreciated when a larger set of indicators is analysed, involving more issues and different fields of concern. For instance, in 1997 Italy was 18.3 years ahead of Slovenia for GDP per capita at purchasing power parity, while Slovenia was 1.6 years ahead of Italy for Internet hosts per capita. Some of these indicators can change very quickly, some others, like some demographic variables and some other characteristics of human factor, very slowly. Time distances will be different, smaller for those indicators that are more dynamic by their nature, more conducive to policy measures and given higher priority in decision-making process.

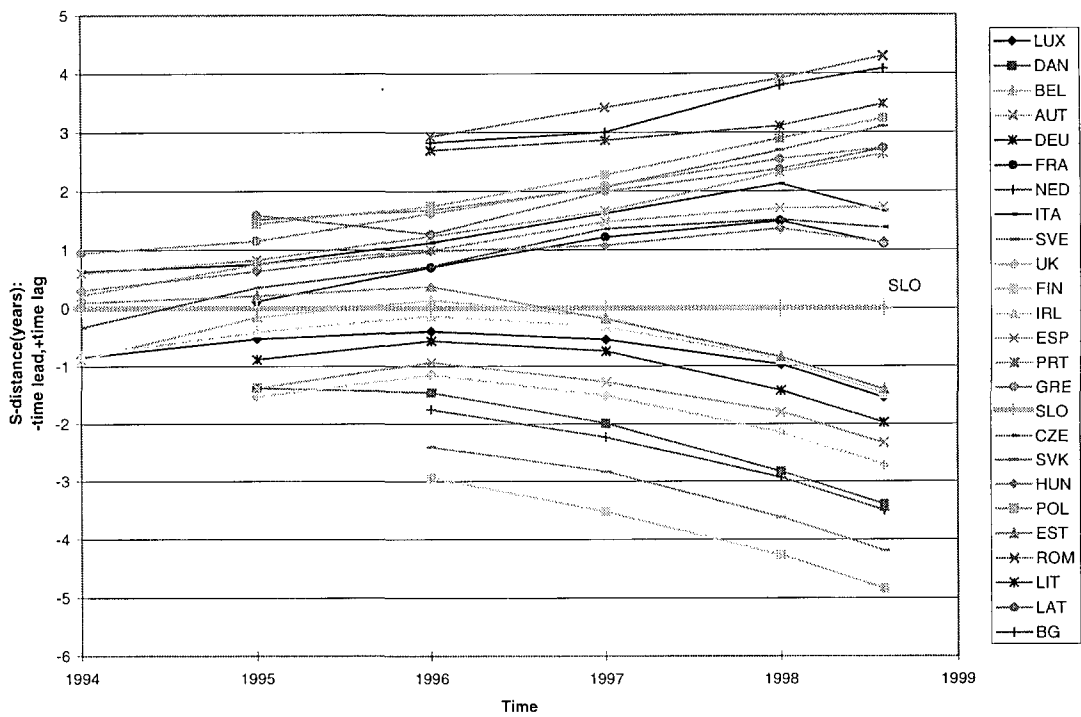


Figure 1. Time distance for Internet host density per 10000 inhabitants, EU and candidate countries, Slovenia=0

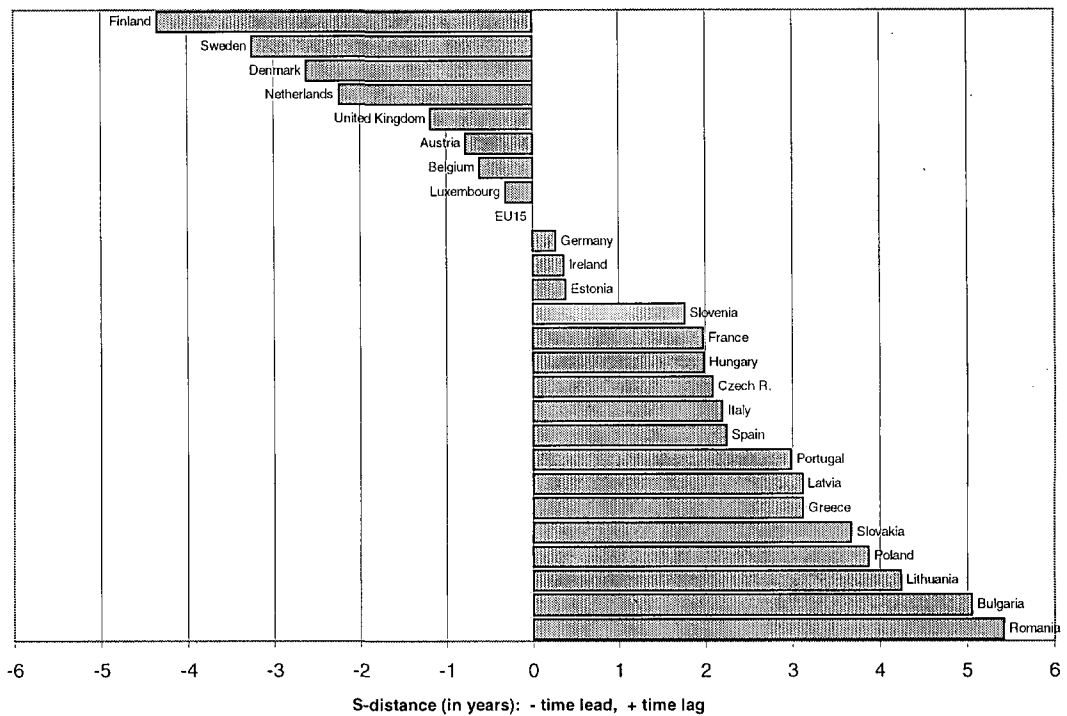


Figure 2. Differences from EU15 average for Internet hosts per capita expressed in time (August 1999)

Figure 2 is an illustration of application of time distance presentation in a similar case of comparative analysis. In this example the average value of Internet hosts per capita for EU15 is the benchmark for comparison. The dispersion of situations in this respect for EU15 countries and Central and Eastern European candidate countries can be presented in various ways, like ratios, percentages, absolute value and absolute differences, etc. Furthermore, various summary measures of dispersion could be calculated.

Absolute values of the indicator are presented in Table 2. A widely used conventional measure would be indices or percentage differences. For instance, in August 1999 the index for forerunner Finland would be 467, for Portugal and Greece about 33 as the lowest value for EU15 countries, and 9 for Bulgaria and 7 for Romania (EU15=100). Figure 2 presents another complementary view of this set of data. Time distances are calculated in cases of above the average countries for the level of EU15 average, and in cases of below the average countries for the level of indicator in these countries in August 1999. Finland had a lead of about 4.5 years ahead of the EU15 average, Portugal and Greece were lagging the EU15 average for about 3 years, and Bulgaria and Romania for more than 4 years. Time distances allow for a distinct new insight that can help to form a richer perception of the situation.

Since time distance is expressed in units of time, which everybody understands from ministers, managers to general public, it possesses one of the ideal characteristics of a presentation and communication instrument. It is expected that the analysis of and discussion about time distances will have considerable influence on how people will form their perception about a situation and on public opinion. For instance, in the EU the consideration of economic and social cohesion is an important goal. A series of presentation of results like Figure 2 for a number of relevant indicators would without any doubt provide a new additional insight to a complex multidimensional problem. Similarly, it would be very useful if the results in Table 3 and in Figures 1 and 2 would be provided for a broad selection of information society indicators.

This offers improved semantics for analysis and policy debate, and can in many cases lead to qualitatively different conclusions from those reached in a static conceptual and analytical framework. By analogy, there is a wide-open possibility to apply this methodology to numerous business problems at the micro, corporate and sector levels. Another important advantage of this approach is that the results and conclusions based on the two-dimensional analysis add new information and new insight, while none of the earlier results are lost or replaced.

4 Conclusions

In empirical research the art of handling and understanding of different views of data is crucial for discovering the relevant patterns. The time distance approach (with associated statistical measure S-distance) is useful at least in two domains: it offers a new view of data that is exceptionally easy to understand and communicate, and it may allow for developing and exploring new hypotheses and perspectives that cannot be adequately dealt without the new concept.

The generic nature of the time distance concept and the S-distance measure leads to the conclusion that the methodology can be usefully applied as an important analytical and presentation tool in numerous applications in a wide variety of substantive fields. Especially in the field of information technology indicators, which is characterised by great speed of change, it would be of great interest to complement rather than replace the conventional measurement of differences between countries or other units with this new perspective of the situation.

5 References

- [1] P. Sicherl. *A Novel Methodology for Comparisons in Time and Space*. East European Series No. 45. Vienna Institute for Advanced Studies. Vienna. 1997.
- [2] P. Sicherl. Time Distance in Economics and Statistics; Concept, Statistical Measure and Examples. In A. Ferligoj (ed.). *Advances in Methodology, Data Analysis, and Statistics*. Metodoloski zvezki. 14. FDV. Ljubljana. 1998.
- [3] P. Sicherl. The Time Dimension of Disparities in the World. *XIIIth World Congress of the International Economic Association*. Buenos Aires. August 1999.
- [4] P. Sicherl. Measuring disparities in two dimensions: proximity in time and proximity in indicator space. *10th International Conference on Socio-economics*, Vienna, Austria 13-16 July 1998.
- [5] International Telecommunication Union. *Database*. Geneva. 1998.
- [6] <http://www.ripe.net>
- [7] <http://www.ris.org>
- [8] National Science Board, *Science & Engineering Indicators - 1998*. Arlington, VA: National Science Foundation, 1998.
- [9] V. Vehovar, Spremljanje informacijske družbe, in P. Sicherl, A. Vahcic (eds.), *Model indikatorjev za podporo odlocanju o razvojni politiki in za spremljanje izvajanja SGRS*, Sicenter, Ljubljana, oktober 1999.