Analysis of Information Society Indicators with Time Distance Methodology

Pavle Sicherl

SICENTER (Socio-economic Indicators Center) and University of Ljubljana, Slovenia

A novel generic statistical measure S-time-distance complementing existing methods of analysis of time series data is briefly presented. The application to indicator analysis shows that the gap between compared units may be very different when compared with commonly used static measures and with time distance measure, leading to a special typology of indicators. Analysis of indicators like PC per 100 inhabitants and Internet users per capita will show the empirical results for EU countries.

Keywords: time distance, S-time-distance, information society indicators.

1. Introduction

Time is, besides money, one of the most important reference frameworks in a modern society. Yet the present state-of- the-art does not fully utilize the information content with regard to certain aspects of the time dimension in the existing data. The new generic time distance approach (with associated novel statistical measure S-time-distance) offers a new view of data that is exceptionally easy to understand and communicate, and it allows for developing and exploring new hypotheses and perspectives.

This approach is applied to present a novel typology of indicators and to show the results of the time distance range for 25 EU countries, USA and Japan for two information society indicators.

2. S-Time-Distance: Concept and Definition

The art of handling different views of data is crucial for discovering the relevant patterns and for providing a broader framework for policy analysis. The present state-of-the-art does not realize that, in addition to static comparison, there exists in principle a theoretically equally universal measure of difference (distance) in time when a given level of the variable is attained by the two compared time series. In graphical terms, the usual way is to compare the time series in the **vertical dimension**, i.e. for a given point in time. The time distance approach uses an additional perspective; it compares the respective time series in the **horizontal dimension**, i.e. for a given level of the variable (see e.g. Sicherl, 1969, 1973, 2004a, 2004b and 2004c).

In the analysis of time series the idea of time distance is a generic concept like static difference and the growth rate over time. Time has been until now used in comparisons mainly as location information, i.e. as a coordinate in a parameter frame forming a coordinate system that is used to organize (or index) a set of variables. In other words, it has played the role of a descriptor, subscript or identifier. The new approach offers new avenues for detecting additional information content, without replacing the existing views. If we choose to interchange in the database the roles of the level of the variable and time, a given level of the variable becomes a descriptor or identifier and time becomes a numeraire in which certain distances between the compared units and time series can be expressed and measured (Sicherl, 1997).

Table 1 in Section 4 presents an example with the **time matrix with new information** from which new generic measures can be derived. It shows the time **when a specified level of the variable was achieved in each compared unit**.

Sicherl (2004c) shows that **two operators applied to such time matrix lead to the derivation of two novel statistical measures** expressed in standardized units of time that everybody understands.

The first suggested statistical measure **S-time-distance** measures the distance (proximity) in time between the points in time when the two compared series reach a specified level of the variable X. It compares two series by subtracting **horizontally** the respective times for a given level in the time matrix.

S-time-distance for a given level of XL is defined as[1]

$$(1) \qquad S_{ij}(X_L) = \Delta t(X_L) = t_i(X_L) - t_j(X_L)$$

The sign of the time distance comparing two units (i, j) is important to distinguish whether we are dealing with time lead $(-)$ or time lag $(+)$ (in a statistical sense and not as a functional relationship)

$$(2) \qquad\qquad S_{ij}(X_L) = -S_{ji}(X_L).$$

S-time-distance is calculated from the original values of the variable (with some possible interpolation and extrapolation) without referring to any other information than levels of the variable and time subscripts. This is a confirmation of the statement that time distance provides additional dimensions of description of the state of a multidimensional space of n variables $(X_i, \ i = 1, \dots n)$.

Subtracting the respective times in the time matrix for consecutive levels of the variable for each column **vertically** derives the second suggested measure **S-time-step**. These vertical differences can be labeled as time steps and represent an alternative description to the growth rate measure. This second statistical measure S-time-step and its relation to S-time-distance will not be discussed further here.

Beyond the brief example in this paper, this generic approach can be usefully applied as an important analytical and presentation tool to a wide variety of substantive fields at macro and micro levels. For extensions to measuring deviations between estimated and actual values in regressions and models, forecasting, error in timing and causality, monitoring and business cycle analysis see Sicherl (1994, 1997),

to variables other than time see Sicherl (1999). Granger and Jeon (1997, 2003a) extended it to comparisons of leading and lagging indicators and used the time distance as a criterion for evaluating forecasting models. They also analysed four models of inflation in the USA, not only with the standard method of average squared deviations between the projected and actual values, but also with the time distance method deviations, which produced significantly different results (Granger and Jeon, 2003b).

3. Typology of Indicators

Theoretically, the higher the rate of growth of the indicator, ceteris paribus, the smaller the respective time distance. Empirically, the degree of disparities may be very different in static terms and in time distance, providing new insights from existing data. The present state-of-the-art neglects this additional information that has been always available in time series databases as "a hidden dimension" and thus leads to an information loss that has no justification.

Many information society indicators describe the high growth phenomena in this field. When compared with many economic and social indicators, the growth dynamics in this field is considerably higher. The fact that different indicators exhibit very different growth rates leads within this approach to a typology of indicators based on the relation between static measure(s) and time distance. Comparing 29 OECD countries with the benchmark for three representative indicators clearly shows that, with respect to the three representative indicators, very different conclusions about the magnitude of the gap are reached if one uses static measure of the gap (in this case index) or S-time-distance.

Indicators can be classified into Type I (low static disparity and large time distances), Type II (large static disparities and small time distances), and Type III indicators (for details see Sicherl, 2004a). In Figure 1 the example for Type I is female life expectancy, for Type II Internet users per capita, and for Type III GDP per capita. It is important to emphasize that both dimensions of the gap should be studied simultaneously to obtain a realistic perception of the situation.

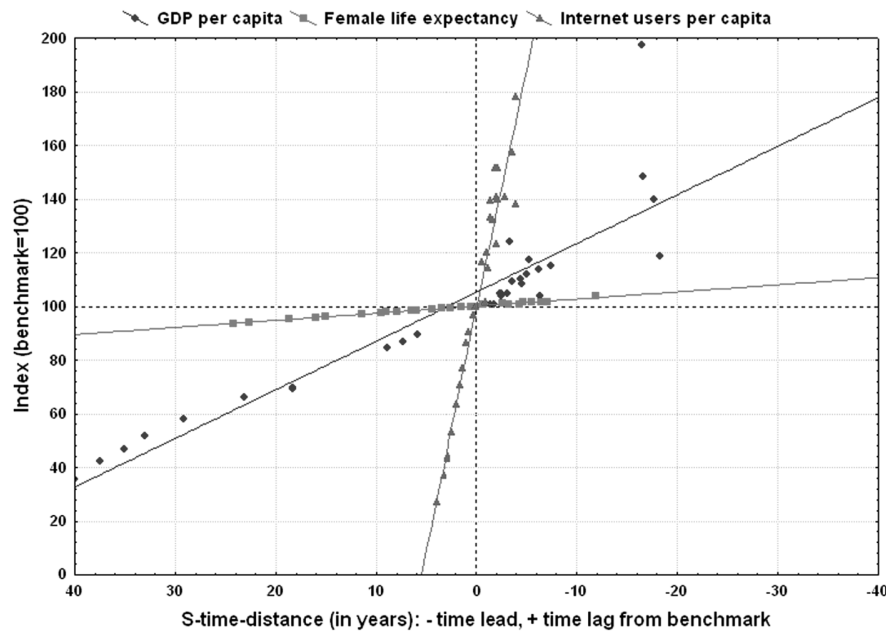[1] For details see Sicherl (2002), also on possible multiple time intersections.

Fig. 1. Evaluation of the degree of disparity may be very different in static terms and in time distance (two-dimensional presentation adds new insights from existing data in a dynamic framework) Source: Sicherl (2004e), data relate to 2002.

Such a classification helps to broaden the horizon and introduces new insights and semantics for analysis and policy debate. Obviously, due to wide differences in growth rates across different fields of concern in a dynamic world, it is implausible to rely only on static measures of disparity. They take into account only differences in growth rates between the units, but are insensitive to absolute magnitudes of growth rates between indicators in the same or different fields of concern.

4. Results for PC per 100 Inhabitants and Internet Users per Capita

In this section we shall first look at the indicator PC per 100 inhabitants in Table 1 to present an empirical example for the time matrix mentioned in Section 2. It is an application of the generic idea that databases can be analysed also by levels of the indicator as the focus of attention on which the time distance methodology is based. For personal computers per 100 inhabitants levels of penetration rates in steps of 5% were arbitrarily selected and by the interpolation of original time series Eurostat data for the period 1990 − 2001 the respective times were calculated. The advantage of such a time matrix table is its graphical quality of presentation, providing a number of observations to a search-

ing mind. It has table-graph combination qualities. It is sometimes very difficult to observe details in a trend graph when you have 18 units in the figure. Not all possible comparisons from such a table-graph will be mentioned here, but only a few. First, one immediately sees which levels were reached by the analysed countries. Second, one also grasps over how many level classes they have advanced in the time span of the period of analysis. Third, for a given level of the indicator one could read off the *S*-time-distance value for that level: e.g. for the level of 40% Sweden was in January 1999 one year and three months behind the USA and two years ahead of Denmark, etc. but the time distance with EU15 at that level cannot be determined since EU15 average has not reached that level yet (Sicherl, 2003).

A more simple presentation of time distances from a benchmark at a given year is presented in Figures 2 and 3. Here the benchmark is EU15 average. Since we can see from Table 1 that dispersion among the analysed countries is high, it would be difficult to find intersections for only one given level of the indicator. Therefore, we use an approximation. For the countries with higher values in 2002 than EU15 average the level of comparison is that of the benchmark unit and the negative values indicate the time lead of these countries.

Level(%)	5	10	15	20	25	30	35	40	45	50	55	60
EU15			Apr-96	Mar-98	Dec-99	Aug-01						
BE			Jul-94	Jul-96	Feb-98	Jul-99	Jun-01					
DK				Feb-95	Sep-95	Nov-96	Oct-97	Aug-99				
DE			Dec-94	Sep-96	Apr-98	Jan-00	Oct-01					
EL	Sep-98											
ES	Jan-95	Apr-98	Mar-01									
FR			Mar-96	Feb-98	Jun-99	Nov-00						
IE			Sep-94	Aug-96	Apr-98	Aug-99	Oct-00					
IT		May-97	Sep-99									
LU							Jan-95	Jan-00	Nov-00			
NL			Mar-94	Dec-95	May-97	May-98	Aug-99	Jan-01				
AT		Dec-93	Oct-95	Sep-97	Aug-99							
PT	Jul-95	Jun-00	Apr-01	Oct-01								
FI			Jun-94	Jul-95	May-96	Sep-97	Dec-98	Feb-01				
SE			Dec-93	Mar-96	Aug-96	Feb-97	Mar-98	Jan-99	Dec-99	Nov-00	Oct-01	
UK				Dec-95	May-98	Nov-99	May-01					
US						Jan-95	Jul-96	Oct-97	Oct-98	Sep-99	Jun-00	May-01
JP		Apr-95	Sep-96	Nov-97	Mar-99	Jun-00						

Table 1. Time matrix: time when a given indicator level was attained (penetration rates for PC per 100 inhabitants in per cent).

For the countries with lower values of the indicator, their respective level of the indicator is the level for which time distance is estimated. This means that time distance indicates the lag of these countries behind the benchmark, i.e. how many years earlier had the EU15 average attained the 2002 level of the indicator for a given country. Comparing Figures 2 and 3 we see that time distances are shorter for Internet users than for PC users, because the former grow faster.

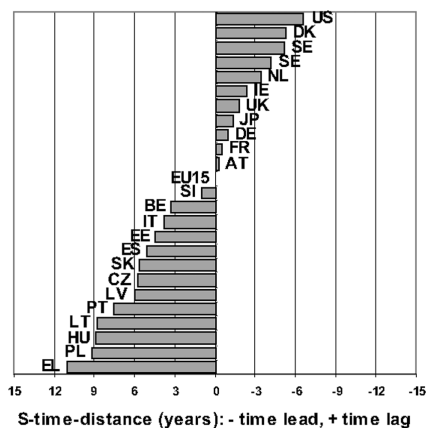A further example shows the application of time distance methodology to the digital divide between various socio-economic groups within the EU (Sicherl, 2004d). The data for this estimation originate from the SIBIS project (SIBIS, 2003) and are based on Eurobarometer surveys in January 1997, October 2000 and SIBIS survey of April 2002. *S*-time-distances in the figure show how many months earlier were the April 2002 values of Internet usage for the selected socio-economic and demographic groups achieved by the average Internet usage for EU15.

Fig. 3. Time distance between EU15 average and selected countries: Internet users per capita, 2002.

Fig. 2. Time distance between EU15 average and selected countries: personal computers per 100 inhabitants, 2002.

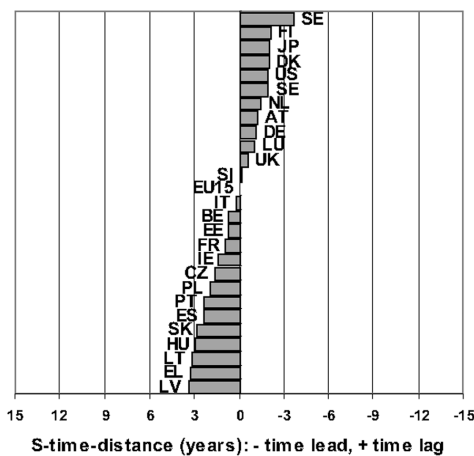The smallest time lag is that for gender, followed by age (+50), income (1stQ) and low education. While the gender time lag for 'total Internet usage' behind the average usage is only about 5 months, for the low education group it is more than 4 years. Furthermore, one can compare penetration rates for different indicators and different categories, which may be of interest also for market analysis. The time lag for 'total Internet usage at home' behind 'total Internet usage' was about 8 months, for some groups it is slightly longer.
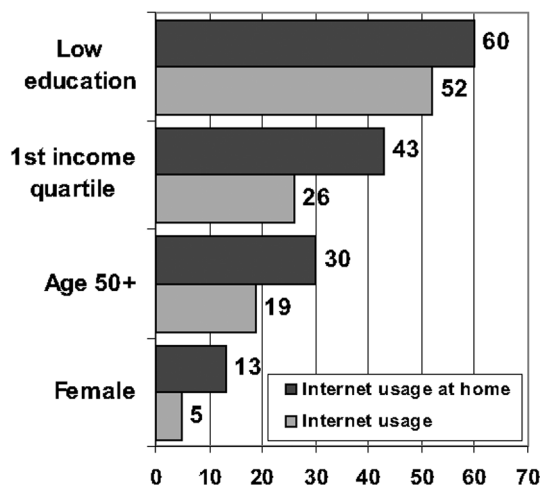

Fig. 4. S-time-distance in months:time lag behind the average Internet usage (base=0).

5. Conclusions

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 in statistical and comparative analysis. Expressed in time units, it is an excellent presentation tool easily understood by policy makers, managers, media and general public and it can support decision-making and influence public opinion. In an information age a new view of the existing databases should be evaluated as an important contribution towards a more efficient utilisation of the available information complementing, rather than substituting, the existing methods in extracting the relevant information content and new insights from available data.

Time distance approach brings about two persuasive advantages for extensive use. First, expressed in time units, *S*-time-distance is comparable across variables, fields of concern, and units of comparison. Second, earlier results are left unchanged, but new conclusions may be reached due to an added dimension of analysis.

Comparing representatives of information society indicators with some economic and social indicators it was shown that wide differences in the average growth rates of indicators in different fields lead to a novel typology of indicators. The gaps between different units should be evaluated both in static measures and time distances simultaneously. Empirical results show that new insight from existing data can be obtained by complementing other methods with the time distance approach and the novel generic *S*-time-distance measure.

References

[1] C. W. J. GRANGER, Y. JEON, (1997) Measuring Lag Structure in Forecasting Models — The Introduction of Time Distance. *Discussion Paper* 97–24, University of California, San Diego.

[2] C. W. J. GRANGER, Y. JEON, (2003a). A Time-distance Criterion for Evaluating Forecasting Models. *International Journal of Forecasting* 19, pp. 199–215.

[3] C. W. J. GRANGER, Y. JEON, (2003b). Comparing forecasts of inflation using time distance, *International Journal of Forecasting*, Vol. 19.

[4] SIBIS, (2003). Measuring the Information Society in the EU, the EU Accession Countries, Switzerland and the US. Empirica, Bonn.

[5] P. SICHERL, (1969). Analiza nekih elemenata za ocenu stepena razvijenosti republika i pokrajina, *Ekonomska analiza* št. 1–2.

[6] P. SICHERL, (1973). Time Distance as a Dynamic Measure of Disparities in Social and Economic Development. *Kyklos XXVI*, Fasc. 3, pp. 559–575.

[7] P. SICHERL, (1994). Time Distance as an Additional Measure of Discrepancy between Actual and Estimated Values in Time Series Models. *International Symposium on Economic Modelling*. The World Bank, Washington D.C.

[8] P. SICHERL, (1997). Time Distance Measure in Economic Modelling. Paper presented at the *International Symposium on Economic Modelling*. University of London, July 23–25, London.

[9] P. SICHERL, (1999). A New View in Comparative Analysis. *IB Revija XXXIII*, pp. 22–34.

[10] P. SICHERL, (2002). The Time Distance among Se-
lected EU and Candidate Countries. *10th General
Conference of European Association of Develop-
ment Institutes*, 19–21 September, Ljubljana.

[11] P. SICHERL, (2003). Different Statistical Measures
Provide Different Perspectives on Digital Divide.
6th conference of ESA, September 23, Murcia.

[12] P. SICHERL, (2004a). Time-distance Analysis:
Method and Applications. eWISDOM 2a/2004,
collection of articles in a thematic issue on time
distance, pp. 1–99.

[13] P. SICHERL, (2004b). Comparing in Two Dimen-
sions: A Broader Concept and a Novel Statistical
Measure of the Time Dimension of Disparities.
European Societies, 6(2), pp. 181–203.

[14] P. SICHERL, (2004c). Time distance: a missing link
in comparative analysis. *Paper for the 28th General
Conference of the International Association for Re-
search in Income and Wealth*, August 22–28, 2004
Cork, Ireland.

[15] P. SICHERL, (2004d) A New Generic Statistical Mea-
sure in Dynamic Gap Analysis. In European Com-
mission, *The European e-Business Report*, 2004
edition, Luxembourg.

[16] P. SICHERL, (2004e). Time distance as a new ad-
ditional way to measure and assess the overall
position among and within countries. *Working pa-
per, SICENTER*, Ljubljana.

Contact address:
Pavle Sicherl
Socio-economic Indicators Center and University of Ljubljana
Brajnikova 19
Ljubljana
Slovenia
`pavle.sicherl@sicenter.si`

PAVLE SICHERL is the founder of SICENTER and Professor of Eco-
nomics at the Faculty of Law, University of Ljubljana, Slovenia. His
specialitys are growth and inequality. He has introduced a new statisti-
cal measure — S-time-distance, to amend existing methods of analysing
disparities in many fields. He was a visiting fellow at Yale, LSE and
other universities. He has also acted as consultant in international de-
velopment matters to the World Bank, OECD, UN, ILO, UNIDO and
Harvard Institute.