Advances in Methodology, Data Analysis, and Statistics
Anuška Ferligoj (Editor), Metodološki zvezki, 14
Ljubljana: FDV, 1998

Time Distance in Economics and Statistics
Concept, Statistical Measure and Examples

Pavle Sicherl[1]

Abstract

A novel statistical measure S-distance (expressed in standardised units - time) is generalised to complement conventional measures in time series comparisons, regressions, models, forecasting and monitoring, and to provide from existing data new insights due to an added dimension of analysis.

1 Introduction

In decision making process comparisons play an important role. The better the analytical framework the greater the information content provided to experts and decision makers. The present state-of-art of comparative analysis, over many dimensions and over time, needs improvement at least in three directions: (1) comparisons over space and over time need to be better integrated, (2) explicit treatment of the time dimension as a universal unit of measurement can contribute new insights to the problem under consideration, and (3) the information content of existing data can be better exploited in a dynamic conceptual and analytical framework.

Time and money are used as two most important common units of measurement (in addition to the physical units relevant for the problem under consideration) to assess and compare various situations. It is remarkable, however, that the present methods in economics and statistics do not fully utilize the information content with regard to certain aspects of the time dimension embodied in existing data. The fact that from practically the same information - two vectors of values with time subscripts - an additional theoretically universal and practically relevant measure can be obtained, is a clear indication that the information content of available data could be used more efficiently.

[1] Department of Economics, Law Faculty, University of Ljubljana and SICENTER, Kongresni trg 12, 1000 Ljubljana, Slovenia

2 Definition of S-distance

Time distance in general means the difference in time between the points in time or periods when the compared events occurred. Sicherl (1973, 1978) defined a special category of time distance related to the level of the analysed indicator: S-distance measures the distance (proximity) in time when the two compared series achieve a given level of the indicator. For an indicator X S-distance is defined as the distance in time (the number of years, quarters, months, weeks, days, etc.) between the points in time when the two series compared attain a specified level of the indicator X. The observed distance in time is used as a dynamic (temporal) measure of disparity between the two series in the same way that the observed difference (absolute or relative) at a given point in time is used as a static measure of disparity.

If the development of the indicator X over time is expressed as $X_i = f_i(t)$ and $X_j = f_j(t)$, or simply $X_i(t)$ and $X_j(t)$, the quantitative estimates of the static and dynamic measures of disparity between the two compared units (i) and (j) are obtained in the following way:

1. When the two functions are compared vertically at a given point of time (t), the static dimension of the disparity is observed. Two of the most frequently used quantitative measures of the static dimension are the absolute static difference A

$$A_{ij(t)} = X_{i(t)} - X_{j(t)} \tag{1}$$

and the relative static difference R

$$R_{ij(t)} = X_{i(t)}/X_{j(t)} \tag{2}$$

2. When the two functions are compared horizontally (i.e. for a given level of the indicator X), the difference represents the time distance between the two units for that level of X. For a given level X_L,

$$X_L = X_i(t_i) = X_j(t_j) \tag{3}$$

and the S-distance separating unit (i) and unit (j) for the level X_L will be written as

$$S_{ij}(X_L) = \Delta T(X_L) = t_i(X_L) - t_j(X_L) \tag{4}$$

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 X (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 ($T_1, T_2, \ldots, T_n$).

The sign of the time distance comparing two units is important to distinguish whether it is a time lead or time lag (in statistical sense and not as a functional relationship)[2]

$$S_{ij}(X_L) = -S_{ji}(X_L) \qquad (5)$$

For a given level of the indicator X_L in general there will be two vectors of the values of time when this level of the indicator (or its approximation by interpolation or extrapolation) will be attained by unit (i) and unit (j): $T_i(X_L)$ with m values and $T_j(X_L)$ with n values. The corresponding matrix of time distances will have m times n elements. In the case of continuously increasing or decreasing series there will be only one time distance. For the strengths and weaknesses of time distance measure and its relationship with the conventional static measures see Sicherl (1997a).

A shorter more convenient notation can be used which places the needed subscripts in brackets:

$$S(CBB(t)) = \Delta T(B(t)) = TC(B(t))-TB(B(t)) \qquad (6)$$

where S-distance S(CBB(t)) represents time distances $S_{CBB(t)}$, i.e. time distances between the compared unit (C) and the base unit (B) calculated for the levels of base unit $X_{B(t)}$. This will probably be the most frequently used specification of time distance in practical applications. However, the more general specification of time distance is related to the level of the indicator X_L which is not necessarily related to a given point in time. In such a general case the simplified notation used will be:

$$S(ijL) = S_{ij}(X_L) = \Delta T(X_L) = T_i(X_L)-T_j(X_L). \qquad (7)$$

[2] The earlier definition of S-distance (see e.g. Sicherl, 1973, 1978 and 1992) used positive sign for time lead and negative sign for time lag. With the generalisation of application of the time distance concept and S-distance measure to short term economic analysis (e.g. deviations in regressions, models, forecasting and monitoring) in Sicherl (1994), for a more clear two-dimensional graphical presentation of deviations between actual and estimated values it was found to be more convenient to assign the negative sign to time lead and positive sign to time lag, as it is implied in equation (4).

3 Proximity in time and proximity in space comparing two units

As in time series applications all events are dated in time, the time distance was always there as a "hidden" dimension. What was needed was to systemise and formalise the approach for operational use once the generalisation from the broader conceptual framework was achieved. This has been done on two levels: conceptual and analytical. The conventional approach does not realise that in addition to the disparity (difference, distance) in the indicator space at a given point in time (e.g. between country A and country B), 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. From this idea of the multidimensional notion of disparity (proximity) it follows that the overall degree of disparity (proximity) is conceived here as a weighted combination of the static and temporal dimensions of disparity.

In practical applications it is important to distinguish backward looking (*ex post*) and forward looking (*ex ante*) time distances. They relate to different periods, past and future. Backward looking time distance belongs to the domain of statistical measures based on known facts, and there is no need to relate it to any static measure or growth rate. The second, the forward looking time distance, is important for describing the time distance outcomes of alternative assumptions about future developments or of alternative policy scenarios for the future.

The examples presented below will offer illustrations of empirical application of time distance analysis from two broad fields. The first will be the field of long term analysis where it will be illustrated in a development and wellfare context. The second example will be application of time distance as a measure of discrepancy and goodness-of-fit in time series regressions, models and forecasting in short term analysis.

Sicherl (1978) elaborated the conceptual and methodological approach for development and welfare issues. Empirical applications between two or more units, like regional comparisons in the former Yugoslavia (Sicherl, 1992), gender disparities (Sicherl, 1989), comparisons across neighbouring countries (Sicherl, 1990), disparities among EU countries and regions (Sicherl, 1997a), and infant mortality as an example of studying social welfare across Europe (Rose, 1992), provided interesting new insights for the problem under scrutiny in this domain.

One example from long-term analysis should suffice to illustrate the possibility that the qualitative conclusion based on this methodology can be very different from that arrived at by conventional analysis. In the field of distribution of income an interesting example is the disparity between median income for white and black families in the USA where the conclusions about the degree of disparity are quite different using the suggested approach from those based on conventional static measures alone. Table 1 shows the values of median family income for white and black families in 1990 CPI-U adjusted dollars. Bureau of the Census provides data for black families starting only in 1967. Median white family income was increasing until 1973, while in the period 1973-1990 both the median white family income and median black family income remained approximately constant (with considerable short-term fluctuations).

The relative disparity between white and black family income has not changed very much, for the period 1967-1990 black family income amounted to about 60 per cent of white family income. Also the absolute difference between the family income of the two groups stayed approximately constant. The conventional analysis would thus conclude that the degree of disparity between the two compared groups remained unchanged.

Time distance analysis shows a distinctly different picture. Time distance between the two groups was about 17 years in 1970 and about 35 years in 1990. Obviously, when looked upon from both dimensions, the position of black families in 1990 is further away from the comparable position of white families than in 1970. The time distance estimate used in this example is historical backward looking time distance, showing that the absolute level of median income of black families in 1990 was achieved by white families around 1953, i.e. about 35 years ago. This lag in 1990 is twice the time lag that existed in 1970.

In policy-oriented research three types of issues are involved: (1) estimation of statistical measures of disparities, which can be thought of as 'objective' measures of the multidimensional notion of distances in space and time for many indicators, (2) value judgements associated with them, which give subjective weights to the 'objective' measures across various dimensions and fields of concern, (3) analysis of behaviour as reactions of people to the level and change in the extent of disparity (Sicherl, 1992).

An interesting comment was provided by an American professor: 'George Bush should have known it before Los Angeles riots'. This broader conceptual and analytical framework would not tell him when and where such a situation could take place, but the perception of the change in the degree of inequality would be qualitatively different. Namely, static measure(s) of disparity, based on US Bureau of Census (1991) data, would indicate that there was no substantial change of inequality between the income of white and black families in the USA between 1970 and 1990. Time distance between the two groups for this indicator, however, was about 17 years in 1970 and about 35 years in 1990. The broader concept, where proximity in time is one dimension in a multidimensional concept of disparity, would warn policy makers that the feeling of inequality in the society has increased (at least in one dimension), whereas the conventional relative static measure would not communicate such a message. This example illustrates how the time distance approach deals with some weaknesses of existing concepts and methods (Sicherl, 1996b).

Table 1: Static difference and time distance between median income for white and black families, USA 1967-1990 (in 1990 CPI-U adjusted dollars)

Year	Median family income		Black income as	Time distance	Abs. difference
	White	Black	percent of white	S(BWB) in years	A (B-W)
1947	18503				
1948	17951				
1949	17749				
1950	18683				
1951	19399				
1952	20291				
1953	21529				
1954	21077				
1955	22497				
1956	24035				
1957	24038				
1958	23969				
1959	25345				
1960	25765				
1961	26144				
1962	26993				
1963	27968				
1964	28914				
1965	30086				
1966	31566				
1967	32221	19077	59.2	16.5	-13144
1968	33565	20131	60.0	16.2	-13434
1969	34879	21364	61.3	16.1	-13515
1970	34481	21151	61.3	17.3	-13330
1971	34440	20783	60.3	18.6	-13657
1972	36111	21462	59.4	19.1	-14649
1973	37076	21398	57.7	20.1	-15678
1974	35546	21225	59.7	19.9	-14321
1975	34662	21327	61.5	20.8	-13335
1976	35689	21229	59.5	21.9	-14460
1977	36104	20625	57.1	24.7	-15479
1978	36821	21808	59.2	23.5	-15013
1979	36796	20836	56.6	26.6	-15960
1980	34743	20103	57.9	28.2	-14640
1981	33814	19074	56.4	30.5	-14740
1982	33322	18417	55.3	32.3	-14905
1983	33905	19108	56.4	32.4	-14797
1984	34827	19411	55.7	33.0	-15416
1985	35410	20390	57.6	32.9	-15020
1986	36740	20993	57.1	33.4	-15747
1987	37260	20177	54.2	35.1	-17083
1988	37470	21355	57.0	33.8	-16115
1989	37919	21301	56.2	34.8	-16618
1990	36915	21423	58.0	35.8	-15492

Data source: US Bureau of Census (1991), p.201

4 S-distance as a measure of discrepancy between actual and estimated values

Sicherl (1994) introduced a new direction of the generalisation of the time distance approach, which includes the extension and application of this approach to the measurement of discrepancy between the estimated and actual values in time series regressions and models, forecasting and monitoring. The change from the earlier case of comparisons between different units in Section 3 comes in these applications in Section 4 from the choice of what we wish to describe, analyse and compare. Here the comparison is between two different states (positions) of the same unit for the analysed indicator, i.e. between the actual value and the estimated (forecast, budgeted, planned, targeted, etc.) value.

Statistical measure S-distance is defined in such a way that all results (and the underlying assumptions) of time series regressions and models are left unchanged, and the time distance analysis is added at the end of each stage to explore the additional information content that can be provided by time distance analysis. Such a position makes it possible to specify a general procedure for calculating S-distance between estimated and actual values S(CBB(t)) for multiple regressions and model estimates. B(t) represents a time series of actual values, and C(t) a time series of comparator series, in this case estimated values. The estimated relation is C(t)=f(Z(t)). Z(t) can be a vector (time for time trend or some other variable like capital, employment, interest rate, etc.) or a matrix (e.g. in multiple regressions with more independent variables).

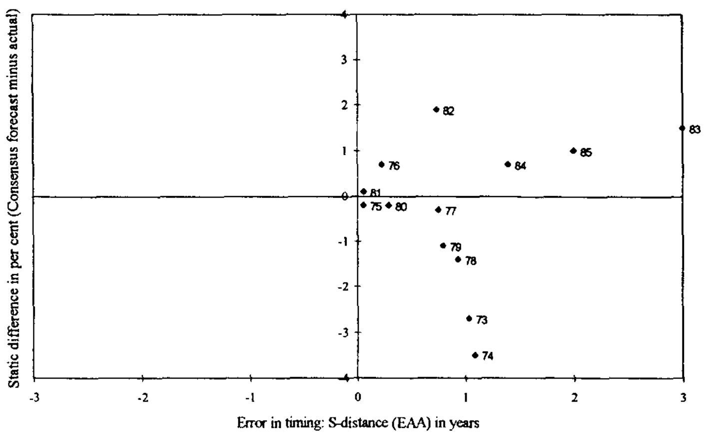

Figure 1: Two-dimensional presentation of errors (static difference in growth rate of the deflator and time distance) of Consensus forecast for GNP deflator, USA (1973-1985)

For both regressions and model simulation it is in general possible to obtain two vectors of values of dependent variable(s) needed to calculate S-distance: time series of actual values B(t) and time series of estimated values C(t). For such a case an alternative description of time distance S(CBB(t)) can be written as S(EAA(t)), indicating in more specific terms that this is S-distance between estimated and actual values, calculated for the level of actual value(s) of the dependent variable.

The second example relates to evaluation of forecasting procedures and signals that are important for improvements in forecasting and error correction. It is obvious that inflation rate in the USA is very important for many decisions, including the financial markets. The Consensus forecast as the mean value of forecasts of the most important USA forecasting institutions, including major banks and corporations, is evidently based on a vast accumulated knowledge and resources. Each point in Figure 1 shows the deviation of Consensus forecast from actual values of USA inflation rate for a given year in the analysed period 1973-1985 (data source: Artis, 1988) in two dimensions. On the horizontal axis the error in timing is represented by S-distance S(EAA), whereas on the vertical axis the conventional deviation in inflation rate at a given point in time is represented.

The two-dimensional presentation of deviations of estimated values from the actual values can now exhibit four theoretically possible deviations from the actual value in the observed period: the estimates in the first quadrant are too high and too late, in the second too high and too early, in the third too low and too early and in the fourth quadrant too low and too late. Beside S-distances between actual and estimated values for individual points in the time series, by analogy with the conventional standard error of the estimate (SEE) in the indicator space it is in principle possible (for well-behaved series) to calculate a summary measure of the goodness-of-fit with respect to timing - standard error in time (SET). Such a complementary measure would be of interest for each case separately, but also in comparing different cases (Sicherl, 1994).

The conventional methodology would find the Consensus forecasts in Figure 1 unbiased as far as high and low estimates are concerned. However, this methodology finds Consensus forecasts on USA inflation for the analysed period biased as they are practically always too late for a given level of the indicator. The fact that for a given level of actual inflation rate the consensus forecast was practically always too late leads to a different evaluation of the forecasting results. Since the actual series is from this novel perspective (for a given level of the inflation rate) a 'leading indicator' for forecast values, these forecast cannot be considered satisfactory, either from the statistical or logical standpoint. This example gives an idea of the potential of this methodology to provide new insights from the existing data for a variety of situations that are too numerous to be enumerated.[3]

[3] By analogy this methodology could be applied to numerous similar problems in business at the micro and corporate levels (e.g. in monitoring by comparing production and financial targets with actual implementation).

5 S-distance analysis for a single time series

S-distance is a time distance that is defined for a given level of the indicator. It should be emphasised that in principle time distance is independent of the static distance(s) for a given point in time. Two time series (e.g. actual and estimated values of the analysed indicator X) can be analysed independently from two perspectives: for a given point in time one gets static deviations, for a given level time distances as deviation in timing. There are certain advantages and disadvantages of the two approaches.

The advantage of comparisons at a given point in time lies in the fact that, for the known length of the two time series, all values can be determined as a single value for any given point in time. For comparisons at a given level of the indicator, time distance cannot be determined (without extrapolation or interpolation) for those levels of the indicator which were not reached by both series. Also, it is possible that for the same level of the indicator there are multiple crossings and the analysis of time distances becomes more complex. On the other hand, if the two time series are continuously increasing or decreasing, when defined also for time distances one gets single values only for a given level of the indicator.

In general, while the levels or static differences can be written as a function of time, time intersections and time distances for a given level of the variable have to be expressed as relations. Thus in the computer programme for calculation of S-distances developed by SICENTER one has the options for calculating them on the basis of the first or last intersection, as an average of all time distances or as a minimum distance. As always, it is for the user to decide which is the most useful for the purpose of the inquiry (Sicherl, 1997b).

In business cycle analysis and more generally in analysing stationary series with considerable fluctuations, for the same chosen level of the indicator there could be multiple crossings and the analysis of time distance for the selected level of the indicator becomes more complex. The third group of S-distance applications relates to the calculation of the indicator for a single time series, which make no sense in static comparison, but can be a useful device to study some characteristics of time series with considerable fluctuations.

There are multiple intersections of the curve exhibiting the movement of the analysed indicator X over time with the selected level of the indicator X_L, represented by vector T_1, T_2, ..., T_n. The corresponding matrix of S-distances for the chosen level of the indicator X_L is obtained formally in the same way as in the case when one compares two different series, though in this case the matrix is symmetric and the elements above the diagonal differ from the corresponding elements below the diagonal only in sign. The most interesting S-distances from such a matrix are S-distances between the neighbouring intersections in time.

For this type of S-distances it may be convenient to use a slightly different notation. If using the previous notation S-distances for a single time series for a given level of the indicator X_L would be described as S(BBL). Since there are multiple crossings and possibly numerous S-distances, we shall for the case of S-distances for a single time series use the notation $S_{i(i,j,L)}$, where S_i denotes S-distance for the series on itself (i.e. for a single time series), while the first and the second position in the brackets indicate the sequential index of the given crossing between the indicator curve and the level X_L for which the S-distance is defined and calculated. Time distances associated with neighbouring intersections are thus $S_i(_{i,i+1,L})$, where each of them is linked with a corresponding value T_i. Such S-distances fall into two groups: those for which the indicator curve lies above or below the time interval for which S-distance $S_{i(i,i+1,L)}$ is defined (Sicherl 1996a).

With such information one can study the pattern of time distances at any level of the indicator. If one uses the example of business cycle analysis, determination of time distances for various phases of the cycle need not be focused only on peaks and troughs as it is usually done in business cycle analysis. Also, since the amplitude of peaks and troughs are not the same in different cycles, the more flexible and detailed analysis offered by this methodology could be a useful complementary tool to the present state-of-the-art. Another example of possible aplication of S-distance analysis for a single time series is time distance analysis of the time series of the error term from regressions and models. Needless to say, there is a broad selection of possible field of aplication, variables and respective statistical measures which could profit from this new view of data.

6 Conclusions

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 a wide variety of substantive fields. One can by and large expect the benefits derived from an additional descriptive and presentation measure offering a fresh perspective on the situation under scrutiny in all time series applications. Even if this would be the only benefit of its use, it would be unwise not to take advantage of a new analytical tool. However, in some cases, like examples of the evaluation of the degree of inequality by race in the USA, or the evaluation of the Consensus forecast of the inflation rate in the USA, the broader conceptual and analytical framework also qualitatively changed the 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.

References

[1] Artis, M.F. (1988): How Accurate Is the World Economic Outlook? A Post Mortem on Short-Term Forecasting at the International Monetary Fund, *Staff Studies*, Washington, IMF.

[2] Rose, R. (1992): Making Progress and Catching Up, A Time-Space Analysis of Social Welfare Across Europe, *Studies in Public Policy*, **208**, Glasgow, CSPP.

[3] Sicherl, P. (1973): Time-Distance as a Dnamic Measure of Disparties in Social and Economic Development, *Kyklos*, **XXVI, 3**, 559-575.

[4] Sicherl, P. (1978): S-distance as a Measure of the Time Dimension of Disparities. In: Z. Mlinar and H. Teune, (Eds.): *The Social Ecology of Change*, Beverly Hills/London: Sage, 237-268.

[5] Sicherl, P. (1989): *Methods of Measuring Disparity Between Men and Women*, Santo Domingo: United Nations International Research and Training Institute for the Advancement of Women.

[6] Sicherl, P. (1990): *Slovenija zdaj* (Slovenia Now), Ljubljana: Gospodarski vestnik.

[7] Sicherl, P. (1992): Integrating Comparisons Across Time and Space: Methodology and Application to Disparities within Yugoslavia, *Journal of Public Policy*, **12, 4**, 377-403.

[8] Sicherl, P. (1994): Time Distance as an Additional Measure of Discrepancy Between Actual and Estimated Values in Time Series Models, Paper presented at the International Symposium on Economic Modelling, Washington, World Bank, June 22-24 1994.

[9] Sicherl, P. (1996a): Time Distance in Business Cycle Analysis: Methodology and Examples for the USA, Japan and Germany, International Symposium on Economic Modelling, Statistics Oslo, Norway, 3-5 July, 1996.

[10] Sicherl, P. (1996b): Perceiving and Measuring Disparities in Time: An Integration of Static and Intertemporal Comparisons, Paper for the 24th General Conference of the International Association for Research in Income and Wealth, Lillehammer, 18-24 August 1996.

[11] Sicherl, P. (1997a): Time Distance Measure in Economic Modelling (Paper presented at the International Symposium on Economic Modelling, University of London, London, July 23-25, 1997), Ljubljana: Sicenter, May.

[12] Sicherl, P. (1997b): *A Novel Methodology for Comparisons in Time and Space*, East European Series No. 45, Vienna: Institute for Advanced Studies, 1997.

[13] US Bureau of Census (1991): Money Income of Households, Families, and Persons in the United States: 1990, *Current Population Reports, Series P-60, No. 174*, Washington.