Volume Title: The Interpolation of Time Series by Related Series

Volume Author/Editor: Milton Friedman

Volume Publisher: NBER

Volume ISBN: 0-87014-422-7

Volume URL: http://www.nber.org/books/frie62-1

Publication Date: 1962

Chapter Title: The Form in Which to Use the Data

Chapter Author: Milton Friedman

Chapter URL: http://www.nber.org/chapters/c2066

Chapter pages in book: (p. 22 - 23)

Of course, if we take the assumptions strictly, and so suppose that nothing at all can be said about the value of μ, this result cannot help us. But it may well be that enough is known to indicate whether the coefficient of variation of v is smaller or larger than unity. If it is smaller than unity, M_β', will be better than M_1, and conversely.

IV. THE FORM IN WHICH TO USE THE DATA

Each method of interpolation we have been discussing is itself a set of methods, depending on the form in which the original data are expressed—whether as arithmetic observations, logarithms of the original observations, ratios of the observations to arithmetic straight-line trends connecting values for dates at which the series to be interpolated is known, etc.

I shall not attempt to explore systematically the choice of the form in which to express the data. Rather, I shall simply list the considerations suggested by the preceding analysis that are relevant to the choice. The form should, if possible, be chosen to satisfy three conditions: (1) to assure that $\mu_u = \mu_v = 0$; (2) to render the series of values of u and v for different dates homogeneous; and (3) to facilitate the accurate estimation of the required parameters.

The primary means of satisfying condition (1) is through the choice of the trend values to be associated with each unknown value of the series to be interpolated and with the corresponding value of the interpolator. The selection of the proper trend value is precisely the problem of mathematical interpolation without the aid of related series. One requirement likely to be imposed on mathematical interpolation is that it yield an unbiased estimate, which is identical with the satisfaction of condition (1). Mathematical interpolation may therefore be regarded as a first step, yielding as a first approximation what we have called the trend value, to be improved by the use of a related series. In practice, for interpolation of monthly intervals (or other time intervals shorter than a year), what is here called the trend value includes the seasonal component.

The deviation from the trend can then be computed so as to satisfy condition (2). In general, the chief problem here will be to make the standard deviation the same for different dates. For economic data, it is generally supposed that the coefficient of variation is more likely to be the same over time than the standard deviation, which suggests a logarithmic or relative transformation.

The same transformation of the data may be used to compute the trend value and the deviation from trend as, for example, when logarithms of the original data are used throughout. But this need not be done. For example, relatives to arithmetical straight-line trends involve two different transformations. Let x_i' be the observations in the form in which they come. Then the use of relatives to trend is equivalent to the transformation

$$x_i = \frac{x_i'}{(1 - w_i)x_0' + w_i x_2'}, \tag{43}$$

so the data are combined arithmetically in computing trend values, after which relatives are used to compute deviations from trend. Similarly, instead of (43), one could use

$$x_i = \log x_i' - \log [(1 - w_i)x_0' + w_i x_2'], \tag{44}$$

which combines logarithmic and arithmetic transformations.

Condition (3)—facilitating the accurate estimation of the required parameters—becomes a somewhat independent condition when the parameters are estimated from "test" series. It then dominates the transformations to be applied to these series. This problem is only touched on in the present paper and needs much further attention.

V. CONCLUSIONS

This paper deals with the problem of estimating intermediate values of a time series. This can be done by mathematical interpolation using only the known values of the given time series or by using one or more related series whose values are known for the desired time intervals and whose movements are supposed to be correlated with the movements of the series to be interpolated. These notes are restricted to the simplest form of the problem: that in which only the known values immediately preceding and following the value to be interpolated are used explicitly in mathematical interpolation and in which only one related series is used. This simple case probably covers the great bulk of the interpolation performed in practice.

The major conclusions of our analysis can be summarized as follows:

1. Mathematical interpolation and interpolation by related series are not substitute methods; rather they are complementary. In the words of our practical maxim I: *First interpolate mathematically.* In general, getting as good an estimate by mathematical interpolation as possible will involve transforming the data into a form in which straight-line interpolation can be expected to give unbiased estimates of the unknown values or of seasonally adjusted unknown values if interpolation is done for intervals that are fractions of a year. In the words of our practical maxim II: *Carry out interpolation by related series with seasonally adjusted data.* If the final series is desired in seasonally unadjusted form, the seasonal should be estimated and combined with the values obtained by straight-line interpolation, even if the seasonal is estimated from the related series. This gives a first approximation to the unknown values. Call it the "trend value."

2. A related series can then be used to improve this approximation by providing an estimate of the deviation of the unknown value from it.

3. For this purpose, trend values of the related series should be obtained by mathematical interpolation as in point (1), including the seasonal component, if any.

4. The related series and its trend values should be expressed in a form (logarithmic, relative to trend, etc.) for which the deviations of the transformed series from the similarly transformed trend values can be expected to be homogeneous over time and linearly related to the deviations, correspondingly transformed, of the original series.