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: Conclusions

Chapter Author: Milton Friedman

Chapter URL: http://www.nber.org/chapters/c2067

Chapter pages in book: (p. 23 - 29)

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.

5. An estimate should be made from "test" series or otherwise of the standard deviation of these deviations for the related series (σ_v) and also for the series to be interpolated (σ_u) and of the correlation coefficient of the two sets of deviations (ρ_{uv}).

6. The size of the correlation coefficient determines the extent of improvement that can be attained by use of the related series. Some improvement is possible as long as it is not zero.

7. Compute an estimate of $\beta = \rho_{uv}\sigma_u/\sigma_v$ from the estimates in point (5). Call this estimate $\hat{b}$. Take $\hat{b}$ times the deviation of the related series as an estimate of the deviation of the series being interpolated. This is method $M_{\hat{b}}$.

8. Methods widely used transfer the deviation of the related series in full to the series being interpolated. This is equivalent to using a value of $\hat{b} = 1$, hence is designated method M_1. Method M_1 will worsen rather than improve matters unless $\rho_{uv}\sigma_u/\sigma_v$ is greater than 1/2 and will seldom yield smaller errors on the average than $M_{\hat{b}}$. It may nonetheless be worth using on some occasions simply because it requires less information and hence is less costly.

9. In practice, M_1 has been used so widely and uncritically that it must have often yielded poorer results than mathematical interpolation alone and may well have done so more frequently than it has yielded better results. One reason for such an outcome is that the relevant criterion in choosing related series and in judging the extent of improvement possible through their use is the size of ρ_{uv}. This correlation may be small even though the correlation between the original and the related series at known dates is high, yet the latter is often the criterion implicitly or explicitly used to choose and judge interpolators.

10. A point that is important for practical work arises when only one component of a broader total is unknown for the desired dates. In such a case, in the words of our practical maxim III: *Perform interpolation only on the part of a series that is unknown for the dates for which interpolation is to be done; never on a broader total, part of which is known for those dates.*

APPENDIX NOTES

1. *Relation between parameters of the distributions of the original and transformed variables.* Let μ and σ represent the mean and standard deviation of the variable to be designated by a subscript and ρ the correlation coefficient between the two variables designated by subscripts. Further, consider the more general transformation

$$u_i = x_i - [(1 - w_i)x_0 + w_i x_2] \qquad \text{(i)}$$

$$v_i = y_i - [(1 - w_i)y_0 + w_i y_2], \qquad \text{(ii)}$$

where w_i is the relative weight attached to the terminal value in computing the straight-line trend. This transformation covers explicitly not only the case in the text but also both nonequally spaced intervals and more than one intermediate value to be interpolated. The relation between the parameters of the (u_i, v_i) universe and those of the $(x_0, x_i, x_2, y_0, y_i, y_2)$ universe is then as follows:

$$\mu_{u_i} = \mu_{x_i} - (1 - w_i)\mu_{x_0} - w_i\mu_{x_2} \qquad \text{(iii)}$$

$$\sigma_{u_i}^2 = \sigma_{x_i}^2 + (1 - w_i)^2\sigma_{x_0}^2 + w_i^2\sigma_{x_2}^2 - 2(1 - w_i)\rho_{x_0x_i}\sigma_{x_0}\sigma_{x_i} - 2w_i\rho_{x_ix_2}\sigma_{x_i}\sigma_{x_2} + 2w_i(1 - w_i)\rho_{x_0x_2}\sigma_{x_0}\sigma_{x_2}, \tag{iv}$$

with μ_{v_i} and $\sigma_{v_i}^2$ given by the same formulas except that y replaces x in all subscripts;

$$\begin{aligned}\rho_{u_iv_i}\sigma_{u_i}\sigma_{v_i} = {} & \rho_{x_iy_i}\sigma_{x_i}\sigma_{y_i} + (1 - w_i)^2\rho_{x_0y_0}\sigma_{x_0}\sigma_{y_0} + w_i^2\rho_{x_2y_2}\sigma_{x_2}\sigma_{y_2} \\ & - (1 - w_i)[\rho_{x_iy_0}\sigma_{x_i}\sigma_{y_0} + \rho_{x_0y_i}\sigma_{x_0}\sigma_{y_i}] \\ & - w_i[\rho_{x_iy_2}\sigma_{x_i}\sigma_{y_2} + \rho_{x_2y_i}\sigma_{x_2}\sigma_{y_i}] \\ & + w_i(1 - w_i)[\rho_{x_0y_2}\sigma_{x_0}\sigma_{y_2} + \rho_{x_2y_0}\sigma_{x_2}\sigma_{y_0}].\end{aligned} \tag{v}$$

If we suppose on grounds of symmetry that

$$\sigma_{x_0} = \sigma_{x_i} = \sigma_{x_2} = \sigma_x; \tag{vi}$$

$$\sigma_{y_0} = \sigma_{y_i} = \sigma_{y_2} = \sigma_y; \tag{vii}$$

$$\rho_{x_0y_0} = \rho_{x_iy_i} = \rho_{x_2y_2} = \rho_{xy}, \tag{viii}$$

then (iv) and (v) reduce to:

$$\sigma_{u_i}^2 = 2\sigma_x^2[1 - w_i + w_i^2 - (1 - w_i)\rho_{x_0x_i} - w_i\rho_{x_ix_2} + w_i(1 - w_i)\rho_{x_0x_2}] \tag{ix}$$

and

$$\begin{aligned}\rho_{u_iv_i}\sigma_{u_i}\sigma_{v_i} = {} & \sigma_x\sigma_y[2\rho_{xy}(1 - w_i + w_i^2) - (1 - w_i)(\rho_{x_0y_i} + \rho_{x_iy_0}) \\ & - w_i(\rho_{x_iy_2} + \rho_{x_2y_i}) + w_i(1 - w_i)(\rho_{x_0y_2} + \rho_{x_2y_0})].\end{aligned} \tag{x}$$

These formulas may suggest the simplicity of exposition gained by the simple transformation (7).

2. *Straight-line interpolation in the light of statistical considerations.* The use of straight-line interpolation in the transformation (i) and (ii) is not at all a self-evident step requiring no justification, once the problem is viewed statistically. From that point of view, the question is how to form the best estimate of x_i from x_0 and x_2 alone. The answer depends very much on what is assumed known about the parameters of the multivariate distribution of the x's. For simplicity, accept the symmetry assumption (vi), return to the case in the text of three equally spaced values of X, and for this case adopt the additional symmetry assumption

$$\rho_{x_0x_1} = \rho_{x_1x_2}. \tag{xi}$$

The least squares estimate of x_1 from x_0 and x_2 is then given by

$$x_1^* = \mu_{x_1} + \frac{\rho_{x_0x_1}}{1 + \rho_{x_0x_2}}(x_0 - \mu_{x_0}) + \frac{\rho_{x_0x_1}}{1 + \rho_{x_0x_2}}(x_2 - \mu_{x_2}). \tag{xii}$$

If we suppose that the μ's are unknown but equal, then the best estimate of all

the μ's from x_0 and x_2 alone is the mean of x_0 and x_2. Inserting these estimates for the μ's in (xii) gives

$$\overset{*}{x}_1 = \frac{x_0 + x_2}{2}, \tag{xiii}$$

or the straight-line estimate. Similarly if the means are supposed unequal but linearly related to time, the same result follows.

Suppose, however, that the x's have already been adjusted for any longer period trend or cycle movements so that all the means can be taken as zero. Equation (xii) would then reduce to (xiii) only if

$$\rho_{x_0x_1} = \frac{1}{2}(1 + \rho_{x_0x_2}). \tag{xiv}$$

It is perhaps clear intuitively why, under these highly special assumptions' straight-line interpolation may not be justified. For example, suppose x_0, x_1, and x_2 are independent, uncorrelated observations with zero means. Then x_0 and x_2 provide no information relevant to estimating x_1 and the best estimate of x_1 is zero.

3. *Relation between M_β and multiple regression method of estimation.* William Kruskal has derived the necessary and sufficient conditions for the estimate of x_1 obtained from the multiple regression of x_1 on x_0, x_2, y_0, y_1, and y_2 to be equal to that given by M_β, when all parameters are assumed known. Let ξ stand for an arbitrary one of the five independent variables and cov ( ) for the covariance of the variables in the brackets. Then the necessary and sufficient condition derived by Kruskal is that there exist two numbers A and B not both zero for which

$$A \text{ cov } (\xi, u) = B \text{ cov } (\xi, v),$$

for ξ equal successively to x_0, x_2, y_0, y_1, and y_2. I am greatly indebted to Kruskal for this proof.

A particular set of conditions satisfying this requirement, and which are therefore sufficient but more stringent than necesary, is the following:

$$\sigma_{x_1}\rho_{x_0x_1} = \frac{1}{2}(\sigma_{x_0} + \sigma_{x_2}\rho_{x_0x_2}) \tag{xv}$$

$$\sigma_{x_1}\rho_{x_1x_2} = \frac{1}{2}(\sigma_{x_2} + \sigma_{x_0}\rho_{x_0x_2}) \tag{xvi}$$

$$\sigma_{x_1}\rho_{x_1y_2} = \frac{1}{2}(\sigma_{x_2}\rho_{x_2y_2} + \sigma_{x_0}\rho_{x_0y_2}) \tag{xvii}$$

$$\sigma_{x_1}\rho_{x_1y_0} = \frac{1}{2}(\sigma_{x_0}\rho_{x_0y_0} + \sigma_{x_2}\rho_{x_2y_0}), \tag{xviii}$$

plus four additional conditions obtained from these by replacing the x's throughout by the corresponding y's and the y's by the corresponding x's.

The meaning of these conditions is that they assume that u and v are each uncorrelated with x_0, x_2, y_0, and y_2, hence that the terminal values contain no information relevant to the prediction of u from v.

If we accept the symmetry assumptions (vi), (vii), and (viii), then these conditions reduce to the following conditions on the correlation coefficients:

$$\rho_{x_0x_1} = \rho_{x_1x_2} = \frac{1}{2}(1 + \rho_{x_0x_2}) \tag{xix}$$

$$\rho_{y_0y_1} = \rho_{y_1y_2} = \frac{1}{2}(1 + \rho_{y_0y_2}) \tag{xx}$$

$$\rho_{x_0y_1} = \rho_{x_1y_2} = \frac{1}{2}(\rho_{xy} + \rho_{x_0y_2}) \tag{xxi}$$

$$\rho_{x_1y_0} = \rho_{x_2y_1} = \frac{1}{2}(\rho_{xy} + \rho_{x_2y_0}). \tag{xxii}$$

Condition (xix) is, of course, the same as conditions (xi) and (xiv) since the problem in Appendix Note 2 is a component of the problem considered in this note.

4. *Preliminary observations on the problem of distribution.* Perhaps the simplest form of the problem of distribution that still preserves its essential features is the conversion of annual data to semi-annual data. To render this analogous to the problem of interpolation, we must consider three years for which we have annual totals for the series to be interpolated and semi-annual figures for the related series. The problem is to distribute the annual total of X for the middle year on the basis of the other data. We must use three years to avoid arbitrariness; if data for the end year contain information for the distribution of the total for the middle year, then data for the initial year must contain the same kind of information.

Let the (unknown) semi-annual values of X be given by $x_1, x_2, x_3, x_4, x_5, x_6$, the (known) semi-annual values of Y by $y_1, y_2, y_3, y_4, y_5, y_6$; the annual values of the two series by

$$\xi_1 = x_1 + x_2,$$
$$\xi_3 = x_3 + x_4,$$
$$\xi_5 = x_5 + x_6,$$

$$\eta_1 = y_1 + y_2,$$
$$\eta_3 = y_3 + y_4,$$
$$\eta_5 = y_5 + y_6.$$

The analogue to the correlation method for interpolation is first to get estimates of x_3 and x_4 from ξ_1, ξ_3, and ξ_5 and then to correct these estimates on the basis of the deviations of y_3 and y_4 from similar estimates based on η_1, η_3, η_5.

In order to assure satisfaction of the constraint

$$\xi_3 = x_3 + x_4, \tag{xxiii}$$

it will be simplest to work with

$$\Delta x_{43} = x_4 - x_3, \tag{xxiv}$$

rather than with x_4 and x_3 themselves. Given an estimate of Δx_{43}, the simultaneous solution of (xxiii) and (xxiv) will give the desired estimates.

The transformation of variables analogous to (7) is then

$$\left.\begin{aligned} u &= \Delta x_{43} - \frac{1}{8}(\xi_5 - \xi_1) \\ v &= \Delta y_{43} - \frac{1}{8}(\eta_5 - \eta_1) \end{aligned}\right\}. \qquad \text{(xxv)}$$

The rest of the interpolation analysis then carries over directly to these transformed variables. In particular, of course, the three practical maxims apply in full.[16]

As for interpolation, the justification for using $\frac{1}{8}(\xi_5 - \xi_1)$ as an estimate of Δx_{43} is not self-evident. The relevant analysis is similar to that in Appendix Note 2. For simplicity, let us again make symmetry assumptions that the x's all have the same standard deviation and that the serial correlations of x depend only on the interval between the items correlated and not on which particular items are correlated. Let ρ_i equal the correlation between values of x separated by i time units (i.e., by i half-years). All the information provided by the annual totals on Δx_{43} is clearly contained in

$$\begin{aligned} \Delta\xi_{31} &= \xi_3 - \xi_1 \\ \Delta\xi_{53} &= \xi_5 - \xi_3. \end{aligned} \qquad \text{(xxvi)}$$

The multiple regression of Δx_{43} on $\Delta\xi_{31}$ and $\Delta\xi_{53}$ reduces, under our assumptions, to

$$\Delta x_{43}^{*} = \mu_{\Delta x_{43}} + \frac{\rho_1 - \rho_3}{2 + 2\rho_1 - \rho_3 - 2\rho_4 - \rho_5}[\Delta\xi_{51} - \mu_{\Delta\xi_{51}}]. \qquad \text{(xxvii)}$$

If, by analogy with the interpolation problem, we suppose that the mean differences between successive semi-annual observations are unknown but equal, then the best estimate from the annual observations alone of $\mu_{\Delta x_{43}} = \frac{1}{8}\Delta\xi_{51}$, and of $\mu_{\Delta\xi_{51}} = \Delta\xi_{51}$, so (xxvii) reduces to

$$\Delta x_{43}^{*} = \frac{1}{8}(\xi_5 - \xi_1), \qquad \text{(xxviii)}$$

which is identical with the negative of the second term on the right-hand side of (xxv).

If we suppose that the annual totals have already been adjusted for any

---

[16] For n sub-periods, with $n > 2$, the situation though similar is more complex. Given the constraint equivalent to (xxiii), there are only $n-1$ independent variables but there is no particular way of transforming the n variables into $n-1$ that has the direct appeal of (xxiv). However, it does not much matter which way it is done, as long as both X and Y are treated alike. Perhaps the simplest, and the most like (xxiv), is to use the $n-1$ first differences between contiguous observations.

longer-period trend or cycle movements so that all the means and differences can be taken as zero, then (xxvii) would reduce to (xxviii) only if

$$6\rho_1 - 7\rho_3 + 2\rho_4 + \rho_5 = 2. \qquad \text{(xxvix)}$$

Interestingly enough, if the pattern of the ρ's is the natural extension of (xiv), i.e., each correlation coefficient is the arithmetic average of the preceding and following one,[17] the coefficient of $\Delta\xi_{51}$ in (xxvii) turns out to be $\frac{1}{7}$ independently of the value of ρ_1.

[17] This pattern could not of course continue indefinitely, since it would ultimately lead to correlation coefficients less than -1.