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: Errors of Estimation Associated with Noncorrelation Methods

Chapter Author: Milton Friedman

Chapter URL: http://www.nber.org/chapters/c2064

Chapter pages in book: (p. 9 - 14)

designed to convert Y units into X units; they then transfer to X the full amplitude of Y, as converted into X units; and fourth, none of the methods take account of the degree of correlation between X and Y or of the relative amplitude of variation in X and Y.

It may further be worth noting that, among these methods alone, only methods (1), (2), (3), and (4) are satisfactory on formal grounds. Methods (5b) and (6b) reduce to (2); methods (5a) and (6a) in the form in which they are generally used are technically defective since the results depend on a purely arbitrary decision; the variants of (6a) listed in a footnote that are free from this defect are analogous to (3) in their motivation and seem less appealing than (3).

II. ERRORS OF ESTIMATION ASSOCIATED WITH NONCORRELATION METHODS[6]

It will clearly suffice to confine attention to method (1), designated as M_1, in analyzing the errors associated with these noncorrelation methods. The other acceptable methods simply involve applying method (1) to data expressed in a different form—in logarithms, as ratios to an arithmetic trend, or as differences from a geometric trend. In consequence, results for method (1) can be readily translated into corresponding results for the other methods.

A. *Formal specification of* M_1

It will involve no loss of generality to confine our attention to three equally spaced time units, say t_0, t_1, and t_2, for which the values of X are x_0, x_1, x_2, and the values of Y are y_0, y_1, y_2. The values x_0, x_2, and all three values of Y are known. The problem is to estimate the unknown value of X, x_1, by interpolation.

We may further simplify the analysis by expressing our observations as deviations from the corresponding trend values. This mathematical step corresponds to a practical maxim implicit in the preceding section (maxim I): *First interpolate mathematically*. The deviation of the related series from a correspondingly interpolated value can then be used to adjust the interpolated value so obtained. We shall further simplify by using a simple form of mathematical interpolation, namely linear interpolation. Other forms can either be reduced to the linear form by suitable transformations of the data (see Section IV) or require the use of more information than two known observations.

Designate the deviation of X from its trend value by u and the deviation of Y from its trend value by v. We then have

$$\left.\begin{aligned} u_0 &= 0 & v_0 &= 0 \\ u_1 &= x_1 - (1/2)(x_0 + x_2) & v_1 &= y_1 - (1/2)(y_0 + y_2) \\ u_2 &= 0 & v_2 &= 0. \end{aligned}\right\} \qquad (7)$$

[6] Curiously enough, the problems considered in this and the next section are formally identical with those involved in judging the circumstances under which a government policy designed to be countercyclical will in fact succeed in reducing instability, and in specifying the optimum mganitude of countercyclical action. In consequence, these sections largely parallel my article "The Effects of 'Full Employment Policy' on Economic Stability: A Formal Analysis," published in my *Essays in Positive Economics*, Chicago, 1953, pp. 117–32.

Since we shall be concerned solely with u_1 and v_1, we can drop the subscripts and refer simply to u and v.

Let us continue to use an asterisk to indicate an estimate of an unknown value. Then, in our new notation M_1 provides an estimate as follows:

$$_1u^* = v, \tag{8}$$

where the subscript on the left of u^* indicates the method by which the estimate is made.

Expressing our problem in these terms makes it obvious that it is identical with an ancient and elementary statistical problem: given a pair of correlated variables, to predict the value of one variable from the value of the other variable. In its simplest form, this is the problem of simple correlation or elementary regression, which perhaps makes it obvious that (8) is by no means the estimate of u from v that has the lowest average error and that an estimate with a lower average error can be obtained by taking account of the degree of correlation between u and v and the amount of variation in each. But it may be well to proceed more slowly and postpone these considerations to a later stage, both because of the widespread use of (8), i.e., of one or another of the variants of method (1) discussed above, and because the examination of the estimate described by (8) will give us greater insight when we turn to the problem of reducing the error of estimation.

Moreover, the widespread use of (8) may be justified despite the availability of alternative methods yielding lower errors. The use of such alternatives requires more information than the use of (8), hence is likely to be more costly. It may not be worth the extra cost to reduce the error of estimation, especially when a great many interpolations are to be made involving many different series so that only a few interpolations could use the same additional information.

In referring to the average error of M_1, we are implicitly regarding the (unknown) value of u for a particular date together with the known value of v as a random sample of one pair of observations from some bivariate universe. While we have described t_1 as a *particular* date, we are not really interested in the error made in using (8) for any one date but rather in the "average" error that is made (or, more generally, the distribution of errors) in repeated applications of M_1. As always with time series, there are thorny problems about the meaning of the "universe." We shall bypass these problems and simply suppose that u and v can be regarded as a pair of correlated random variables with means μ_u and μ_v, standard deviations σ_u and σ_v, and correlation coefficient, ρ_{uv}. These are the only parameters of the universe that we shall need for what follows. (They completely describe the universe from which a particular pair u, v is regarded as a sample if the universe is bivariate normal. For most of what follows, it is unnecessary to suppose that this is the case, since complete description is not needed.)

[7] See Appendix Note 1 for a discussion of the relation between the parameters describing the distribution of u and v and those describing the distribution of $x_0, x_1, x_2, y_0, y_1, y_2$; and Appendix Note 2 for the statistical justification for the straight-line interpolation entering into transformation (7).

Similarly x_0, x_1, x_2, y_0, y_1, y_2, of which u and v are functions, are to be regarded as a set of random intercorrelated variables, and each sextet of values, as a sample of one from the corresponding universe.[7]

B. *The Mean Values of u and v*

If the straight-line trend is a satisfactory method of mathematical interpolation and so of transforming the X and Y variables into u and v—or, put differently, if the original observations have been expressed in a form that makes a straight line trend satisfactory—then μ_u and μ_v might be expected in general to equal zero, which is to say that the deviations from the straight-line trend might be expected to average zero.

Perhaps the chief exception relevant to practical work arises when interpolation is for intrayear observations of a series that is subject to a seasonal. If the seasonal of X is known—as it may be from another period or other evidence even if the value of X is not—an obvious improvement over straight-line interpolation is to superimpose the seasonal deviation on the straight-line trend.

Similarly, in applying M_1 when the seasonal in both X and Y are known and are not identical, an obvious improvement is to use the deviation of v from its mean to estimate the deviation of u from its mean, i.e., to use the estimate

$$_1u^* = \mu_u + (v - \mu_v). \tag{8.1}$$

Another possibility is that the seasonal in X is not known but that in Y is and it can be assumed that the two seasonals are identical (i.e., that $\mu_u = \mu_v$). (8) and (8.1) then give the same result. That is, it makes no difference whether a seasonally unadjusted Y series is used to interpolate X by method (1) or a seasonally adjusted series is and then the seasonal in Y is added to X. However, this statement does not hold for the correlation method discussed in the next section, for which it is better to adjust for seasonal before interpolation.

These considerations suggest that, whenever μ_u and μ_v are known, the related series should be used to interpolate the deviation of u from its mean. But obviously this is equivalent to defining u and v in the first place as deviations from the sum of the straight-line trend and the known average of the deviations from the trend, in which case μ_u and μ_v would be equal to zero. In consequence, the only case that needs to be considered under the heading "means of u and v known" is that in which $\mu_u = \mu_v = 0$.

This conclusion can be stated in the form of an important practical maxim (maxim II): *Carry out interpolation with seasonally adjusted data.*[8] If the final series is desired in seasonally unadjusted form, introduce the seasonal after interpolation, and do this even if the seasonal in the interpolated series is estimated from the interpolator.

It is not easy to visualize many practical examples in which it will be desirable to consider μ_u and μ_v as unknown. If the relationship between the two means were also considered unknown, this would be equivalent to ruling v out as an interpolator. A correlation between deviations of u from its mean and

[8] Clearly, this maxim can be regarded as an immediate corollary of maxim I, the particular form of mathematical interpolation being the superposition of a seasonal on a straight-line (or other) trend.

deviations of v from its mean, no matter how high, would be of little use if *nothing* were known about the two means. Consequently, the only case of any interest is that in which something is assumed about the relation between μ_u and μ_v. Such cases can generally be reduced (by changing the units in which v is measured) to the case in which it is assumed that $\mu_u = \mu_v$ but that the common value of the two means is unknown. I believe that in actual interpolation it will seldom be found desirable to use these assumptions. They have, however, considerable theoretical interest since they—and they alone of those so far mentioned—imply that method (1) is in one sense the "best" method of interpolation regardless of the correlation between u and v. We shall postpone detailed consideration of this case until Section IIID. Until that point, we shall deal primarily with the case $\mu_u = \mu_v = 0$. However, we shall derive all relevant formulas in general form.

C. *Comparison of M_1 and Straight-Line Interpolation*

If we use (8), the error for any particular date is the difference between the estimate, u^*, and the correct value, u, or

$$_1d = {}_1u^* - u = v - u. \tag{9}$$

The expected value of the errors—the "bias" in this method of estimation—is

$$\mu_{1d} = E(v - u) = \mu_v - \mu_u, \tag{10}$$

where E stands for expected value. If $\mu_u = \mu_v = 0$, as we shall for the most part suppose, the bias is, of course, zero.

A more important question is the expected value of the error or "average error" in a sense that disregards the sign of the error. It will be convenient to measure this by the mean square error,[9] which will then be given by

$$\begin{aligned} \text{M.S.E. } (M_1) &= E_1d^2 = E(v - u)^2 = \sigma^2_{v-u} + (\mu_{v-u})^2 \\ &= \sigma_v^2 + \sigma_u^2 - 2\rho_{uv}\sigma_u\sigma_v + \mu_v^2 + \mu_u^2 - 2\mu_u\mu_v. \end{aligned} \tag{11}$$

In order to judge whether the mean square error given by (11) is large or small, we may compare it with the mean square error of mathematical interpolation of X, i.e., the mean square error of estimating x_1 as equal to $\frac{1}{2}(x_0+x_2)$ plus a seasonal deviation, if any, or of setting u equal to zero. There is clearly no point to using the related series Y unless doing so reduces the error below that of mathematical interpolation. The gain from using Y can therefore be judged by the fraction of the total error in the mathematically interpolated value that can be eliminated thereby. Accordingly, we shall use straight-line interpolation as our yardstick—let us call it method (0) or M_0. Intuitively, one is likely to expect M_1 to be better than M_0 if there is any positive correlation between the movements of the related series and the series to be interpolated. As we shall see, however, this is not so; if there is positive correlation, it is always possible to do better than M_0, but not necessarily by using M_1.

M_0 can be described as giving an estimate

$$_0u^* = 0. \tag{12}$$

[9] Note that unless the bias is zero, i.e., $\mu_v = \mu_u$, this is *not* the same thing as the variance of the errors.

The error for any particular date in using this estimate is

$$_0d = {}_0u^* - u = -u, \tag{13}$$

so the mean error, or "bias," is

$$\mu_{0d} = -\mu_u, \tag{14}$$

which will, of course, be zero if $\mu_u = 0$; and the mean square error is

$$\text{M.S.E. } (M_0) = E(_0d^2) = E(u^2) = \sigma_u^2 + \mu_u^2. \tag{15}$$

To compare the errors of the two methods, divide (11) by (15), which gives

$$\begin{aligned} \frac{\text{M.S.E. } (M_1)}{\text{M.S.E. } (M_0)} &= \frac{\sigma_v^2 + \sigma_u^2 - 2\rho_{uv}\sigma_u\sigma_v + \mu_v^2 + \mu_u^2 - 2\mu_u\mu_v}{\sigma_u^2 + \mu_u^2} \\ &= 1 + \frac{\sigma_v^2 + \mu_v^2}{\sigma_u^2 + \mu_u^2} - 2\left(\frac{\rho_{uv}\sigma_u\sigma_v + \mu_u\mu_v}{\sigma_u^2 + \mu_u^2}\right). \end{aligned} \tag{16}$$

It follows that the mean square error of M_1 is smaller, the same, or larger than that of M_0 according as

$$\frac{\sigma_v^2 + \mu_v^2}{\sigma_u^2 + \mu_u^2} - 2\left(\frac{\rho_{uv}\sigma_u\sigma_v + \mu_u\mu_v}{\sigma_u^2 + \mu_u^2}\right) \lessgtr 0 \tag{17}$$

or

$$\sigma_v^2 + \mu_v^2 \lessgtr 2(\rho_{uv}\sigma_u\sigma_v + \mu_u\mu_v), \tag{18}$$

or

$$\rho_{uv} \gtrless \frac{1}{2}\frac{\sigma_v}{\sigma_u} + \frac{1}{2}\frac{\mu_v^2 - 2\mu_u\mu_v}{\sigma_u\sigma_v}. \tag{19}$$

For $\mu_v = \mu_u = 0$, this reduces to

$$\rho_{uv} \gtrless \frac{1}{2}\frac{\sigma_v}{\sigma_u}. \tag{20}$$

The fraction by which the mean square error of linear interpolation is reduced by use of method (1) can be derived from (16). When $\mu_u = \mu_v = 0$, it is given by

$$1 - \frac{\text{M.S.E. } (M_1)}{\text{M.S.E. } (M_0)} \doteq \frac{\sigma_v}{\sigma_u}\left(2\rho_{uv} - \frac{\sigma_v}{\sigma_u}\right). \tag{21}$$

If, as is generally the hope, $\sigma_v = \sigma_u$, method (1) will be an improvement over straight-line interpolation if and only if the correlation between u and v exceeds 0.5; if the correlation is less than 0.5, method (1) will lead to larger errors on the average.

This result, which may at first seem surprising, can perhaps be made intuitively plausible by the following considerations. Suppose that the movements of Y were strictly uncorrelated with those of X. Transferring the movements of Y to X would then be equivalent to adding a strictly random series to the straight-line interpolation of X. This would obviously be worse than using the

straight-line interpolations themselves. Now introduce some correlation between the movements in Y and X. Until the correlation reaches $\frac{1}{2}\sigma_v/\sigma_u$, it serves only to offset part of the damage done by the uncorrelated variation in Y; the net effect is an improvement only when the correlation exceeds $\frac{1}{2}\sigma_v/\sigma_u$.

It may be worth emphasizing that the relevant correlation is that between u and v, not that between X and Y. The former may be quite low even though the latter is quite high because of high serial correlation between the successive values of X and Y (see the formulas in Appendix Note 1). This is one of the major reasons why graphic inspection of time series plotted in their original form may be extremely misleading in judging the value of a series as an interpolator. Sad experience persuades me that it is not easy to find an interpolator for which the relevant correlation is above the critical value. In view of the widespread use of method (1), and the rather casual way in which interpolators are often chosen, I would not be at all surprised to find that in practice the use of related series generally gives larger errors than straight-line interpolation.

III. CORRELATION METHODS

A. *Method (b) and the Errors Associated with It*

Both M_1 and M_0 can be regarded as special cases of a more general method, which we may call method (b), or M_b, and which consists of estimating u by the following:

$$_bu^* = bv. \tag{22}$$

If $b=1$, this is M_1; if $b=0$, this is M_0.

The error involved in using M_b is

$$_bd = {}_bu^* - u = bv - u, \tag{23}$$

so the mean error is

$$\mu_{bd} = \mu_{bv-u} = E(bv - u) = b\mu_v - \mu_u, \tag{24}$$

which will, of course, be zero if $\mu_v=\mu_u=0$,[10] and the mean square error is

$$\begin{aligned} \text{M.S.E. } (M_b) &= E({}_bd^2) = E(bv - u)^2 = \sigma^2_{bv-u} + \mu^2_{bv-u} \\ &= \sigma_u^2 + b^2\sigma_v^2 - 2b\rho_{uv}\sigma_u\sigma_v + \mu_u^2 + b^2\mu_v^2 - 2b\mu_u\mu_v. \end{aligned} \tag{25}$$

The results of the preceding section can, of course, all be derived from these formulas by setting b equal to zero and 1, respectively.

We may now ask what value of b (say, β) is optimum in the sense of minimizing the mean square error. Differentiating the right-hand side of (25) and setting the derivative equal to zero gives

$$2\beta\sigma_v^2 - 2\rho_{uv}\sigma_u\sigma_v + 2\beta\mu_v^2 - 2\mu_u\mu_v = 0, \tag{26}$$

[10] But not, as in M_1, if $\mu_v=\mu_u\neq 0$.