PARTIAL TIME REGRESSIONS AS COMPARED WITH INDIVIDUAL TRENDS

By Ragnar Frisch and Frederick V. Waugh[1]

I. INTRODUCTION

There are in common use two methods of handling linear trend in correlation analysis of time series data, first, to base the analysis on *deviations* from trends fitted separately to each original series, and, second to base the analysis on the original series without trend elimination, but instead to introduce *time* itself as one of the variables in a multiple correlation analysis. The first method may be called the individual trend method and the latter the partial time regression method.

There are certain misconceptions about the relative value of the two methods and about the kinds of statistical results that are obtained by the two methods. The following simple example illustrates the situation. Suppose we are studying the relation of sugar consumption to sugar prices. We have data representing total annual consumption and price for several consecutive years. There may be a strong upward trend in consumption due to population increase or to some other factor that changes or shifts the demand curve. If we want to study only the relation of consumption to price, naturally we must eliminate the trend from the consumption before using the data to measure demand elasticity.

But, suppose that, during the period studied, there has been a persistent tendency in sugar prices to decline and that this has caused a long-time increase in consumption. This fact may be just what we want to express when we speak of the relation of sugar consumption to price. This long-time connection between consumption and price may be even more important for our problem than the short-time connection. If so, we must not, of course, eliminate trends from the variables before proceeding to the analysis. On the contrary, in this case the trend variation in price and the trend variation in consumption must be left in the material and *should be allowed to influence the regression coefficient* between consumption and price.

A common idea seems to be that this can be obtained by using the partial time regression method instead of the individual trend method. This is false. The possibility of determining the long time relation

[1] Mr. Waugh is a fellow of the Social Science Research Council and the present study was made in connection with research in Europe made possible by his fellowship.

considered in the above example does not depend on which of the two methods is used, but on a certain criterion regarding the variability type of the material at hand (the criterion is discussed below). The partial trend regression method can *never*, indeed, achieve anything which the individual trend method cannot, because the two methods lead by definition to *identically* the same results. They differ only in the technique of computation used in order to arrive at the results.

This illustrates one of several misconceptions that exist in this field, but there are also others; the various misunderstandings may be briefly classified into the following three points:

1. The significance of the regression coefficients as determined by the two methods. In particular, there exists a misconception as to the meaning of these coefficients as approximations to the underlying "true" relationship between the variables.
2. The significance of the correlation coefficients as determined by the two methods.
3. The significance of the trends. This last question is of particular interest from the point of view of forecasting.

Before proceeding to a more systematic discussion of these points we shall, by quotations from well known statisticians, illustrate the nature of the misunderstandings.

II. QUOTATIONS

The following are excerpts from an article by Bradford B. Smith "The Error of Eliminating Secular Trends and Seasonal Variation before Correlating Time Series," in the *Journal of the American Statistical Association*, December, 1925. He says:

> Should these two series, dependent and independent, chance to have approximately similar trend or seasonal movements, and should these latter be extracted from the two series prior to their correlation, one might under the name of seasonal and trend extract much of the variation by which the two were related, and thus obscure their true relationship. The unconsidered practice of eliminating trend and seasonal from series prior to their correlation is to be looked upon askance, therefore. It is often a serious error.
>
> The escape from this predicament involves no new theoretical consideration. . . . All that is necessary is to remember that fundamentally a numerical description of passage of time is merely taken to represent the magnitude of the combined effect of otherwise unmeasured factors and then this series of "magnitudes" is treated precisely as any other independent factor. (In Bradford Smith's paper "time" is treated as a factor in a multiple regression equation and partial coefficients of regression are obtained to indicate the relation of the dependent variable to this trend and to the other factors). . . . In following this practice, as often occurs, proper methods go hand in hand with better results. On theoretical considerations, correlation coefficients secured by *simultaneous*, or multiple, correlation methods will be as high or higher, and never less, than those resulting from any possible sequence of *consecutive* elimination of the influ-

ence of independent factors from the dependent, of which current methods of eliminating seasonal variations *before* correlating are an example. In actual trials of the two methods the writer has found that the simultaneous solution for trend and seasonal regression curves and curves for other factors always give markedly higher correlations. . . .

Mordecai Ezekiel in his book *Preisvoraussage bei landwirtschaftlichen Erzeugnissen* published in the series of the Frankfurter Gesellschaft für Konjunkturforschung, 1930, says (page 23):— "Oft wird is notwendig sein, den zusammengesetzten Einflusz dieser Kräfte, die mit dem Kalender variieren, zu messen; . . . allerdings ohne dass man dabei der stets vorhandenen Gefahr unterliegen darf, den Einflusz von zeitlich variienden Faktoren Veränderungen zuzuschreiben, die in Wirklichkeit auf andere Ursachen zurückzerfuhren sind. Aus dieser Grunde wird der Einflusz der zeitlich variierden Faktoren am besten durch eine Bestemmung des Trends in der Relation und nicht durch Feststellung desselben bein jeden eizelnen Faktor gewessen."

III. WHAT IS THE "TRUE" RELATIONSHIP

Proceeding now to a more exact statement of the problem, we must first consider the meaning of a "true" relationship, and in what sense such a "true" relationship can be approximated by various empirical methods.

When comparing the results of different methods in time series analysis one must keep clearly in mind the *object* of the analysis. It must be specified which sort of influence it is desired to eliminate, and which sort to preserve. Unless this is specified it has no meaning to say that a certain method will yield a "truer" relationship than another.

Such an expression has a meaning only if it is referred to a given theoretical framework. An empirically determined relation is "true" if it approximates fairly well a certain well-defined *theoretical* relationship, assumed to represent the nature of the phenomenon studied. There does not seem to be any other way of giving a meaning to the expression "a true relationship." For clearness of statement we must therefore first define the nature of the *a priori* relationship that is taken as the ideal.

Let us express this in mathematical terms. We consider a number of variables $x_0, x_1, \cdots, x_n$, the last of them, i.e., x_n, denoting time. We conceive *a priori* of a relation that expresses x_0 in terms of the other variables, if, for simplicity, the relation is assumed linear, it may be written

$$(3.1) \qquad x_0 = \beta_{01\cdot 2\cdots n}\cdot x_1 + \cdots + \beta_{0n\cdot 12\cdots n-1}\cdot x_n$$

where the β's are constants and the variables are measured from their respective means. Each of the constants in (3.1) is conceived so as to represent an *independent* influence on x_0, for instance, the constant $\beta_{0n\cdot 12\cdots n-1}$ represents the independent influence which time may exert directly on x_0, regardless of the particular way in which the other variables $x_1 \cdots x_{n-1}$ happen to evolve. A theoretical relation postulated *a priori* in this way we may call a *structural* relation. The coefficient $\beta_{0n\cdot 12\cdots n-1}$ in (3.1) we may call the partial, structural coefficient between x_0 and time.

Besides the structural relation (3.1) postulated by theory we shall now consider certain relations obtained by the classical statistical procedures. First we consider the regression of x_0 on the other variables, determined by the usual partial regression method, i.e. the relation

$$x_0 = b_{01\cdot 2\cdots n}\cdot x_1 + \cdots + b_{0n\cdot 12\cdots n-1}\cdot x_n \tag{3.2}$$

where the b's are the usual regression coefficients, $b_{0n\cdot 12\cdots n-1}$ we may call the partial trend coefficient of x_0.

On the other hand we may fit to each variable a straight line trend, i.e., a trend of the form

$$x_i = b_{in}x_n.$$

The coefficients b_{in} we call the individual trend coefficient of x_i $(i=0, 1 \cdots n)$. Further we consider the deviations

$$x'_i = x_i - b_{in}x_n$$

and determine the regression of x'_0 on $x_1' \cdots x'_{n-1}$.

This relation is of the form

$$x_0' = b'_{01\cdot 2\cdots n-1}\cdot x_1' + \cdots + b'_{0,n-1\cdot 1\cdots n-2}\cdot x'_{n-1}. \tag{3.3}$$

The coefficients b' in (3.3) we call the deviation-from-trend coefficients, or shorter the b' coefficients, and the coefficients b in (3.2) we call the time-as-a-variable coefficients, or shorter the b coefficients.

The first question we want to raise is: Will the b' coefficients be significant expressions for the structural coefficients, i.e., for the β coefficients? If not, will the b coefficients be more significant expressions for the β coefficients? On this point the misunderstanding seems to be particularly great. The belief seems to exist—the quotations given indicate this—that if trends are present in the variables $x_1 \cdots x_{n-1}$ the coefficient b' will not measure correctly the β, but the coefficient b will. This is entirely wrong. The correct answer is this: If the material at hand satisfies a certain variability condition, then either the b' or the b may be taken as approximations to the β, while if this variability condition is not fulfilled, neither the b' nor the b will represent the β. The b and the b' are always identical, the two methods only represent different ways of computing the same magnitudes.

The variability condition envisaged is this: Each variable x must contain (beside the linear trend and beside the irregularities that are not systematically connected with x_0) an independent non-linear component. In other words, the deviation of x_i from its linear trend must be *significant* from the point of view of the connection between x_i and x_0. And, furthermore, this fluctuation must not be linearly dependent on the deviations of the other variables from their linear trends.

This means that if the relation between x_o and the other variables is assumed linear, it is *by definition impossible to discriminate* between that part of a linear trend in x_o that is caused by linear trends in the other variables and that part which is caused independently by the flow of time. Either we have to leave both these parts of the linear x_o trend in the data, or we have to eliminate both parts. In the first case the regression coefficients between x_o and the other variables will be influenced amongst other factors by whatever independent linear shift there has been in x_0 over the period studied. In the second case this linear shift is eliminated but at the same time all long-time (linear) relations between x_0 and the other variables are eliminated, so that the regression coefficients between x_0 and the other variables will be determined only by the short time fluctuations.

These criteria are consequences of the considerations developed by one of the present authors in an earlier publication.[2] We shall not give any formal proof of these criteria here. We shall confine ourselves to showing that the coefficients b and b' are identical by definition. This will of course be sufficient to exhibit the fallacy of the belief that anything more can be accomplished by the partial trend regression method than by the method of individual trends.

IV. THE IDENTITY OF THE REGRESSIONS DETERMINED BY THE INDIVIDUAL TREND METHOD AND THE PARTIAL TIME REGRESSION METHOD

We consider first the case of two variables (with time as the third variable). Let there be observed a set of values $x^{(1)}$, $x^{(2)} \cdots$ and $y^{(1)}$, $y^{(2)} \cdots$ of the two variables x and y at the points of time $t=t^{(1)}$, $t^{(2)} \cdots$, and let these values be expressed as deviations from their respective means. That is to say, we have $\sum x=0$, $\sum y=0$, $\sum t=0$, where the summations are extended to all the observations.

The moments of the variables we denote

[2] Ragnar Frisch: "Correlation and Scatter in Statistical Variables," *Nordic Statistical Journal,* August 1929, pp. 36—103.

$$m_{xx} = \Sigma x^2 \qquad m_{yy} = \Sigma y^2 \qquad m_{tt} = \Sigma t^2$$
$$m_{xy} = \Sigma xy \qquad m_{xt} = \Sigma xt \qquad m_{yt} = \Sigma yt. \tag{4.1}$$

Let us first measure the relation of y to x by the individual trend method. If linear trends are fitted separately to x and y by the ordinary least squares procedure, the regression coefficient of x on t is m_{xt}/m_{tt} and the regression coefficient of y on t is m_{yt}/m_{tt}. If the deviations from trend are denoted x' and y' we consequently have

$$x' = x - \frac{m_{xt}}{m_{tt}} \cdot t \tag{4.2}$$

$$y' = y - \frac{m_{yt}}{m_{tt}} \cdot t \tag{4.3}$$

and the regression coefficient of y' on x' is

$$b_{y'x'} = \frac{\Sigma \left(x - \frac{m_{xt}}{m_{tt}} \cdot t \right) \cdot \left(y - \frac{m_{yt}}{m_{tt}} \cdot t \right)}{\Sigma \left(x - \frac{m_{xt}}{m_{tt}} \cdot t \right)^2} \tag{4.4}$$

Expanding this we get

$$b_{y'x'} = \frac{m_{tt}m_{xy} - m_{xt}m_{yt}}{m_{tt}m_{xx} - m_{xt}^2} \cdot \tag{4.5}$$

Now let the data be analysed by the partial time regression method with y as the dependent variable and x and t as the independent variables. The regression coefficients of y on x and t are then found by the solution of the following simultaneous equations:

$$m_{xx}b_{yx \cdot t} + m_{xt}b_{yt \cdot x} = m_{xy}$$
$$m_{xt}b_{yx \cdot t} + m_{tt}b_{yt \cdot x} = m_{ty}. \tag{4.6}$$

Multiplying the second equation by m_{xt}/m_{tt} and subtracting from the first we get

$$m_{xx} - \frac{m_{xt}^2}{m_{tt}} b_{yx \cdot t} = m_{xy} - \frac{m_{xt}m_{yt}}{m_{tt}} \tag{4.7}$$

that is

$$b_{yx \cdot t} = \frac{m_{tt}m_{xy} - m_{xt}m_{yt}}{m_{tt}m_{xx} - m_{xt}^2}. \tag{4.8}$$

Similarly we obtain

$$b_{yt \cdot x} = \frac{m_{xx}m_{yt} - m_{xy}m_{xt}}{m_{tt}m_{xx} - m_{xt}^2}. \tag{4.9}$$

But (4.8) is identically the same as (4.5).

Furthermore, (4.9) is identically the same as the coefficient one would get in *estimating* y by the individual trend method. Indeed, if at the point of time t, the independent variable had the value x, then the estimated value y of the dependent variable would be determined as follows. First one would determine the trend value of y as $m_{yt}/m_{tt}\cdot t$, and to this one would add (4.5) times the estimate of deviation from trend in x, which is equal to $(x-m_{xt}/m_{tt}\cdot t)$. The total estimate of y would consequently be

$$(4.10) \qquad y = \left(x - \frac{m_{xt}}{m_{tt}}\cdot t\right)\cdot\frac{m_{tt}m_{xy} - m_{xt}m_{yt}}{m_{tt}m_{xx} - m_{xt}^2} + \frac{m_{yt}}{m_{tt}}\cdot t$$

which reduces to

$$(4.11) \qquad y = \frac{m_{tt}m_{xy} - m_{xt}m_{yt}}{m_{tt}m_{xx} - m_{xt}^2}\cdot x + \frac{m_{xx}m_{yt} - m_{xy}m_{xt}}{m_{tt}m_{xx} - m_{xt}^2}\cdot t.$$

The coefficient of x in this expression is identical with (4.8) and the coefficient of t is identical with (4.9). Therefore *the regression equations given by the two methods are identical.*

A Numerical Example

These results can easily be checked by setting up a simple problem and solving it by the two methods. The following computations may serve as an example.

Observed data

x	t	y
-5	-3	$+6$
-3	-2	$+5$
-2	-1	$+1$
$+2$	0	$+2$
$+1$	$+1$	-3
$+4$	$+2$	-5
$+3$	$+3$	-6

Moments about the mean

$m_{xx} = 68 \qquad m_{xt} = 41 \qquad m_{xy} = -84$

$m_{tt} = 28 \qquad m_{ty} = -60$

Trend coefficients

$$b_{xt} = \frac{41}{28} = 1.464$$

$$b_{yt} = \frac{-60}{28} = 2.143$$

By the individual trend method we consequently get:

Trend values		*Deviations*	
Of x	Of y	x'	y'
-4.393	$+6.429$	$-\ .607$	$-\ .429$
-2.929	$+4.286$	$-\ .071$	$+\ .714$
-1.464	$+2.143$	$-\ .536$	-1.143
0.000	0.000	$+2.000$	$+2.000$
$+1.464$	-2.143	$-\ .464$	$-\ .857$
$+2.929$	-4.286	$+1.071$	$-\ .714$
$+4.393$	-6.429	-1.393	$+\ .429$

$$b_{y'x'} = \frac{m_{x'y'}}{m_{x'x'}} = \frac{3.858}{7.964} = 0.484.$$

And by the method of partial time regression we get:

$$\begin{aligned} 68\, b_{yx.t} + 41\, b_{yt.x} &= -\ 84 \\ 41\, b_{yx.t} + 28\, b_{yt.x} &= -\ 60 \\ \hline 60.036\, b_{yx.t} + 41\, b_{yt.x} &= -\ 87.857 \\ 7.964\, b_{yx.t} &= 3.857 \\ b_{yx.t} &= 0.484 \quad (\text{The same value as } b_{y'x'}) \end{aligned}$$

Consider now the n variables $x_0 \cdots x_{n-1}$ and let time be an $(n+1)$th variable x_n. Let all the variables be measured from their means so that $\sum x_i = 0$ $(i=0, 1 \cdots n)$ where $\sum$ denotes a summation over all the observations. Let $m_{ij} = \sum x_i x_j$ be the moment of the variable x_i with x_j. The regression of the variable x_k on all the others is the linear equation obtained by replacing the row $m_{k0}, m_{k1} \cdots m_{kn}$ in the moment determinant by $x_0, x_1 \cdots, x_n$ and equating the result to 0. In particular, if x_0 is selected as the dependent variable, the regression will be

$$(4.12) \qquad \begin{vmatrix} x_0 & x_1 & \cdots & x_n \\ m_{01} & m_{11} & \cdots & m_{1n} \\ \cdot & \cdot & \cdots & \cdot \\ m_{n0} & m_{n1} & \cdots & m_{nn} \end{vmatrix} = 0.$$

On the other hand, if we fit to x_i $(i=0, 1 \cdots n-1)$ by least squares a straight line in x_n, we obtain the equation $x_i = m_{in}/m_{nn} \cdot x_n$, so that the deviations from trend will be

$$(4.13) \qquad x_i' = x_i - \frac{m_{in}}{m_{nn}} \cdot x_n.$$

The regression of x_0' on $x'_1 \cdots x'_{n-1}$ is consequently

$$(4.14)\qquad \begin{vmatrix} x'_0 x'_1 & \cdots x'_{n-1} \\ m'_{10} m'_{11} & \cdots m'_{1,n-1} \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ m'_{n-1,0} m'_{n-1,1} & \cdots m'_{m-1,n-1} \end{vmatrix} = 0$$

where $m'_{ij}=\sum x_i' x_j'$ are the moments of the deviations. We proposed first to show that the regression coefficient of x_0 on x_i $(i=1, 2 \cdots n-1)$ in (4.12) is identically the same as the coefficient of x'_0 on x'_i in (4.14).

By virtue of (4.13) the moments of the deviations are expressed in terms of the original moments thus:

$$(4.15)\qquad m'_{ij} = m_{ij} - \frac{m_{in}m_{nj}}{m_{nn}} \qquad (i, j = 0, 1 \cdots n)\cdot$$

Obviously $m'_{ij}=0$ whenever $i=n$ or $j=n$, or both $i=n$ and $j=n$.

We subtract from the row $m_{i0}, m_{i1} \cdots m_{in}$ $(i=1, 2 \cdots n-1)$ in (4.12) m_{in}/m_{nn} times the last row. The resulting expression will be the determinant obtained from (4.12) by replacing all the moments m by m' except those in the last row. In particular, the moments in the last column thus obtained will be zero, except the last element of the last column which is maintained equal to m_{nn}.

In other words (4.12) takes the form

$$(4.16)\qquad \begin{vmatrix} x_0 & x_1 & \cdots x_{n-1} & x_n \\ m'_{10} & m'_{11} & \cdots m'_{1,n-1} & 0 \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot \\ m'_{n-1,0} & m_{n-1,1} & \cdots m'_{n-1,n-1} & 0 \\ m_{n0} & m_{n1} & \cdots m_{n,n-1} & m_{nn} \end{vmatrix} = 0.$$

But the determinant (4.14) may be enlarged to

$$(4.17)\qquad \begin{vmatrix} x_0' & x_1' & \cdots x'_{n-1} & 0 \\ m'_{10} & m'_{11} & \cdots m'_{1,n-1} & 0 \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot \\ m'_{n-1,0} & m'_{n-1,1} & \cdots m'_{n-1,n-1} & 0 \\ c_0 & c_1 & \cdots c_{n-1} & 1 \end{vmatrix} = 0$$

where $c_0, c_1 \cdots c_{n-1}$ is a set of arbitrary numbers, for instance m_{n0}, $m_{n1} \cdots m_{n,n-1}$. This shows that when we write (4.12) and (4.14) in the forms

$$(4.18)\qquad a_0x_0 + a_1x_1 + \cdots + a_nx_n = 0$$

and

$$a_0'x_0' + a_1'x_1' + \cdots + a'_{n-1}x'_{n-1} = 0 \tag{4.19}$$

then the first n coefficients $a_0 \cdots a_{n-1}$ in (4.18) will be proportional to the coefficients $a_0' \cdots a'_{n-1}$ in (4.19). Indeed, the a-coefficients will simply be equal to m_{nn} times the a'-coefficients. We consequently have

$$b_{0i.12\cdots n} = b'_{0i.12\cdots n-1} \qquad (i = 1.2 \cdots n-1) \tag{4.20}$$

where b and b' are the coefficients of (3.2) and (3.3) respectively.

Furthermore it is clear that we can take the regression equation (3.3) as a starting point, insert here the values $x_i' = x_i - b_{in}x_n$ and thus get a regression expressing x_0 in terms of $x_1 \cdots x_n$. This latter regression is identical with the regression (3.2). This follows from (4.20) and from the well known fact that

$$b_{0n} = b_{01.2\cdots n}b_{1n} + b_{02.13\cdots n}b_{2n} + \cdots + b_{0n.12\cdots n-1}b_{nn} (b_{nn} = 1). \tag{4.21}$$

Thus the complete regression equation determined by the partial time regression method and the method of individual trends are identical, (3.2) and (3.3) are only two ways of writing the same equation. This also has a bearing on the meaning of trends as a means of forecasting, this aspect of the problem we shall take up in Section 6.

The way in which we pass from (4.12) to (4.16) is in reality nothing but an application of the Gaussian algorithm for solving the normal equations.[3]

This throws an interesting light on the connection between (4.14) and (4.16) and also on the nature of the Gaussian algorithm itself. It shows that the Gaussian algorithm simply consists in this: First one fits to all the variables, except x_n, a "trend" which is linear in x_n. Then to the deviations from trend in all the remaining variables one fits a new "trend" which is linear in the deviation x'_{n-1}, and so on. In this way one finally gets down to a two-variable problem. And this successive fitting of *individual* trends is identical with the determination of a simultaneous partial regression. *The identity of these two processess is just the basis of the Gaussian algorithm.*

V. THE CORRELATION COEFFICIENTS OBTAINED BY THE TWO METHODS

Another misunderstanding seems to be that the higher correlation coefficients obtained by the partial time regression method than by the individual trend method is an expression of the superiority of the first of

[3] Most of the "new" schemes for solving the normal equations that are developed from time to time are nothing but unessential modifications of the Gaussian algorithm. This applies, for instance, to the Doolittle method.

these two methods. The fact that a comparison of correlation coefficients in this case can lead to a belief in the "superiority" of a method that is by definition identical with the "inferior" method, is a striking example of the perfectly imaginary character of much of the reasoning that is currently based on correlation coefficients.

A coefficient of multiple correlation showing the total relationship between one variable such as prices and two others such as consumption and time, must *by its very definition* always be as high as or higher than the gross correlation between price and consumption expressed as deviations from their respective trends, in other words we always have by definition

$$r^2_{y.xt} \geqq r^2_{x'y'}. \tag{5.1}$$

Indeed it is easy to verify that

$$r^2_{y.xt} = 1 - \frac{\Delta_{xyt}}{\Delta_{xt}} \tag{5.2}$$

$$r^2_{x'y'} = 1 - \frac{\Delta_{xyt}}{\Delta_{xt}\Delta_{yt}} \tag{5.3}$$

where Δ_{xt}, Δ_{yt} and Δ_{xyt} denote the correlation determinants in the sets xt, yt and xyt respectively. These are positive definite determinants lying between 0 and 1. Any of them is equal to 1 when, and only when, all the variables in the set are uncorrelated.[4] The formulae (5.2) and (5.3) show that $r^2_{y\cdot xt}$ is always larger than $r^2_{x'y'}$ *no matter what the data are* (provided only that they consist of real numbers). The only exception is the limiting case $r^2_{y\cdot xt} = r^2_{x'y'}$ which occurs when, and only when, y and t are uncorrelated (disregarding the case where there exists a mathematically exact linear relationship in the set xyt). But that is not all. There is a *great* discrepancy between $r^2_{y\cdot xt}$ and $r^2_{x'y'}$ when and only when nearly all the variation in y can be represented simply by a linear trend. From (5.2) and (5.3) we get indeed

$$\frac{1 - r^2_{y.xt}}{1 - r^2_{x'y'}} = \Delta_{yt} = 1 - r^2_{yt}. \tag{5.4}$$

In other words the only thing that a comparison between $r^2_{y\cdot xt}$ and $r^2_{x'y'}$ can tell us *is whether y has a pronounced trend or not.* The comparison does not tell anything about the nature of this trend: whether it is due to a variation in x or to some other cause. It has no meaning as an indication of a "better result," it does not tell us anything about whether the b coefficients or the b' coefficients are the better approxima-

[4] See for instance "Correlation and Scatter."

tions to the "true" β coefficients (as implied for instance in the quotation from Bradford Smith).

Since the comparison of $r_{y.xt}$ and $r_{x'y'}$ indicates only whether y has a pronounced trend or not, and since this fact will be revealed immediately from the plot of y, the computation of these correlations in the present problem will in our opinion serve no significant purpose. It is only a play with formulae which at best is superfluous, but at worst is dangerous, because it may lead to wholly unwarranted conclusions, such as, for example, the conclusion about the "superiority" of the time regression method.

Some concrete examples may illustrate this. In such studies as those concerned with consumption as a function of prices, price as a function of supply, or acreage as a function of previous price, *it is always possible to obtain high coefficients of multiple correlation by including in the simultaneous analysis such factors as time, population,* etc., whenever the period studied is one of considerable change in one of these factors.

One of the authors of the present article once published[5] the results of certain studies of the relation of vegetable prices to quality, based on several hundred observations which indicated total correlations of about 0.70 between certain qualities and the deviation of prices from the daily quotation on the market. He was criticized in a review by Dr. Bean who claimed that the correlation coefficients were too low to be significant. The critic remarked that he had found from experience that correlations of at least 0.90 were necessary for significant results in price analysis: and well he might, because he was working with annual data of supplies and prices covering from eight to ten observations and using multiple curvilinear correlation methods, including in the analysis such factors as trend. In the case of the study of quality, it would have been an easy matter to raise the coefficient of multiple correlation high enough to suit such a critic, indeed it could have been raised practically as near to unity as one pleased by the simple expedient of including as an independent factor in the analysis the daily quotation for standard quality. This was avoided in order to get a truer statement of the actual relation between quality and price.

We do not wish to object to the various results obtained by Dr. Bean, but we do want to point out the uselessness of comparing results of time series analyses merely on the basis of multiple correlation coefficients.

VI. THE MEANING OF THE TRENDS

In the structural relation (3.1) each variable exerts an *independent* influence on x_0. Hence, if trends are present in $x_1 \cdots x_{n-1}$, this will

[5] F. V. Waugh: *Quality as a Determinant of Vegetable Prices*. New York, 1929.

contribute towards the total trend in x_0. This total trend in x_0 may be looked upon as made up of the indirect trends produced by the trends in $x_1 \cdots x_{n-1}$ and of the partial trend caused by the last term in (3.1). Let us express this mathematically remembering the fundamental distinction between the structural coefficients β and the empirically determined regression coefficients b. If x_i is expressed in the form

$$x_i = x_i' + \beta_{in}x_n \qquad (x_n = \text{time}) \qquad\qquad (i = 1, 2 \cdots n)$$

where β_{in} is a constant and x_i' the deviation from trend in x_i, then (3.1) may be written in the form

$$x_0 = \beta_{01\cdot 2\cdots n}x_1' + \cdots + \beta_{0,n-1\cdot 12\cdots n}x'_{n-1} + \beta_{0n}x_n \tag{6.1}$$

where the coefficient β_{0n} is equal to

$$\beta_{0n} = \beta_{01\cdot 2\cdots n}\beta_{1n} + \beta_{02\cdot 13\cdots n}\beta_{2n} + \cdots + \beta_{0,n-1\cdot 12\cdots n}\beta_{n-1,n} + \beta_{0n\cdot 12\cdots n-1}. \tag{6.2}$$

β_{0n} is the *total* structural trend coefficient for x_0, β_{in} are the *individual* structural trend coefficients for the variables $x_i \cdots x_{n-1}$; the products $\beta_{0i\cdot 12\cdots n}\cdot\beta_{in}$ are the *indirect* trend coefficients in x_0; and $\beta_{0n\cdot 12\cdots n-1}$ the *partial* (and direct) trend coefficient in x_0.

If, in the material at hand, the deviations $x'_1 \cdots x'_{n-1}$ from (linear) trends exhibit systematic, linearly, independent variations, the structural coefficients β in (3.1) may be determined by the data, the empirically determined regression coefficients b may then be taken as approximations to the structural coefficients β with the same subscripts.

Consequently if the coefficients $b_{0i\cdot 12\cdots n}$, and the individual coefficients b_{in} $(i=1\cdot 2 \cdots n)$ $(b_{nn}=1)$, are determined empirically we have (on the assumption that the x'-s deviate significantly from linear trends) representations for all the terms in the right member of (6.2), we may consequently say that we also have determined empirically an approximation to the *total* trend coefficient for x_0. Let b_{0*} be the total trend coefficient thus determined by inserting b for β in the right member of (6.2). In other words, b_{0*} is defined as the composite effect of the trends in $x_1 \cdots x_{n-1}$ and the partial trend of x_0, these latter trends being determined by the usual regression procedure.

We may of course also consider the *individual* trend coefficient for x_0 just as for the other variables. If this individual coefficient is determined by the usual regression procedure, i.e. as the coefficient b_{0n} *it becomes just equal to* b_{0*}. This simply follows from (4.21). And from (4.20) it follows that we get exactly the same determination of b_{0*} whether in the right member of (6.2) we use b_{in}' as approximations to β_{in} or we use b_{in}.

We may express this by saying that the partial trend coefficient for x_0 may be determined either by the partial time regression method directly as the usual coefficient $b_{0n\cdot 12\ \cdots\ n-1}$ or it may be determined as the difference between total trend in x_0 and the trends in x_0 that are ascribed to the influence of the various independent factors, *these latter being determined by the individual trend method.* Also, with respect to the interpretation of the trends, the two methods yield consequently exactly the same results, the only difference being a difference in the technique of computation. The fact here discussed is of course only another aspect of the fact proved in Section 4, namely that the complete regression equations determined by the two methods are identical.

The above conclusion may be illustrated by the case of two variables x and y with time t as a third variable. The partial trend coefficient in y is $b_{yt\cdot x}$, and the trend effect in y computed by the individual trend method is equal to $b_{yt}-b_{xt}.\ b_{y'x'}$, i.e., total trend in y minus indirect trend in y attributed to the fact that y depends on x, this latter dependence being determined by the individual trend method. But according to the general formula (4.21) we just have $b_{yt\cdot x}=b_{yt}-b_{xt}b_{y'x'}$. This checks in the numerical example in Section 4 since $-2.143-(1.464)\cdot(0.484)=-2.852$.

How do these considerations affect the use of trends in forecasting? There are three possible procedures for extrapolating a trend into the future: (1) we can project the total trend observed in x_0, simply using the angular coefficient b_{0n}; (2) we can project the partial trend observed in x_0, namely the trend defined by the coefficient $b_{0n.12\ \ldots\ n-1}$; or (3) we can make what we may call a *composition forecast* of the trend in x_0 by forecasting each of the variables $x_1 \cdots x_{n-1}$ and from these forecasts and the relation between x_0 and $x_1 \cdots x_{n-1}$ and the partial trend in x_0, build up a total forecast of the trend in x_0.

If we are to make such a composition forecast of x_0 we must make individual forecasts of $x_1 \cdots x_{n-1}$ and also of the partial trend in x_0. If each of these forecasts is made by the usual regression methods on the basis of the information contained in the material at hand, *the composition forecast of the trend in x_0 will be exactly the same as the direct mechanical projection of the total trend of x_0.* This follows immediately from the above considerations (see in particular formulae (4.21)). In other words the whole process of determining partial regression coefficients in this case is an entirely unnecessary roundabout calculation. This is another aspect of the identity between partial time regression and the individual trend method.

The only reason for using a composition forecast instead of a mechanical projection of the total observed trend in x_0 is to make it possible to utilize some *other information* regarding the probable future trends of the variable $x_1 \cdots x_{n-1}$. Instead of assuming that all the variables, $x_1 \cdots x_{n-1}$, will continue their growth at the average rate observed in the material, we may, for instance, assume that one of them is going to remain stationary at the level it had at the end of the period studied, or we may assume that its growth rate will become less, etc. If such specific guesses are made about the variables $x_1 \cdots x_{n-1}$, the composition forecast for x_0 using the regression (3.2) will not, of course, be identical with the direct mechanical forecast.

To illustrate the difference between a direct mechanical forecast of the total trend and a composition forecast, let x_0 be the consumption of sugar. If we fit a trend line to sugar consumption over the interval of time for which we have observations, and simply project this line into the future, we are assuming that the trends in population, sugar prices, and other factors influencing sugar consumption, will continue at the same rate that we have observed in the past. In some cases such an assumption is admissable, and in others it is not. For example, prices may have dropped so low during the period studied that further declines are very improbable, so that in forecasting the trend of sugar consumption we must reckon with a probable change in the trend of sugar prices which will, in turn, influence the total trend of consumption. To forecast the trend of consumption in this case, we need to know more than its average rate of increase in the past; we need also to know how much of this increase was due to the decrease in prices, so as to be able to say what the trend would have been had prices remained at a level corresponding to that expected in the future. In other words we must use a composition forecast with more or less plausible guesses regarding the future trends of the variables which influence sugar consumption.

Or, to take another example, if x_0 is the yield of potatoes, and $x_1 \cdots x_{n-1}$ represent weather data such as rainfall and temperature, in many cases we have no reason for expecting a real trend in the weather data, and probably can make no better forecast than to assume that the weather factors will vary around averages based on past observations. In this case the trend in x_0 which we would project would be the partial trend $b_{0n.12 \ldots n-1}$, which as we have shown is the same as the total trend in x_0 minus the trend in x_0 attributed to the factors $x_1 \cdots x_{n-1}$. The projection of this partial trend may, then, be considered as one form of the composition forecast in which we forecast that the various independent factors will not continue their observed rate of growth but will vary around their observed average.

Statistical Seminar, University of Norway