# Experience in the Application of Regression Computing to the Seasonal Adjustment of Statistical Time-Series

*In the Monthly Report for March 1957 an in many respects novel procedure, worked out by the staff of our Department of Research and Statistics for eliminating seasonal fluctuations from time-series of economic data, was described in detail. The purpose of the following article is to report on the experience which has in the meantime been gained in the application of this method. First however let us briefly refer to the problems of seasonal adjustment, and to the essential features of the method worked out in the Bank's Department of Research and Statistics.*

*The individual time-series are affected by seasonal influences (for instance the time of year, the date of the major public holidays, the behaviour of purchasers, the arrangement of holidays and the like) in varying degree. In many series the seasonal movement alters from year to year virtually not at all, or only by degrees, but in others it does so to a considerable extent. Between them there are many shades of difference. The more uniform the seasonal behaviour is, the better the series can be "adjusted for the season"; or, per contra, the more unevenly the seasonal movements emerge, the less will be the value of the results of seasonal adjustment, regardless of what procedure is adopted.*

*The starting point for considering a new seasonal adjustment procedure was the knowledge, gained from practical experience, that between the seasonal fluctuations of numerous statistical time-series and their trend — which on this short-term view also contains the cyclical component — there is no strictly proportional relationship, whereas in most of the older methods a strict proportionality between trend and seasonal component is assumed. By reason of the fact that for purposes of the procedure here described the seasonal movement was defined as a correlated connection between series value and trend value it was possible to use the regression method for seasonal adjustment. In this procedure the regression equations are in all cases linearly arranged, in the usual general form, with a constant (additive) term. This approach permits a multiplicative a n d an additive relation between the trend and the seasonal movement being taken into account simultaneously.*

*Since on the first application of this method there were available, in the case of most time-series, only for a few years figures which after the removal of governmental controls again showed a relatively stable seasonal movement, the — arithmetically simple — linear arrangement of the regression equations was held to be expedient. Subsequently this arrangement was in principle retained, but the procedure was applied for moving periods — a simple method permitting allowance for any non-linear regression relationships which might in the course of time emerge.*

*The trend value used for the computing is obtained through systematic translation of the moving twelve-month average, so that the seasonal movement oscillating about it does not appear distorted. Thereby it is possible to make more accurate statements regarding the seasonal behaviour of the series values, and also to discern smaller variations in the direction of movement.*

*So as to evidence the fact that seasonal adjustment is a matter of ascertaining approximate values there is placed around the seasonally adjusted values, where possible, a dispersion band which indicates the exactitude of the seasonal behaviour of the time-series.*

## 1. General Experience

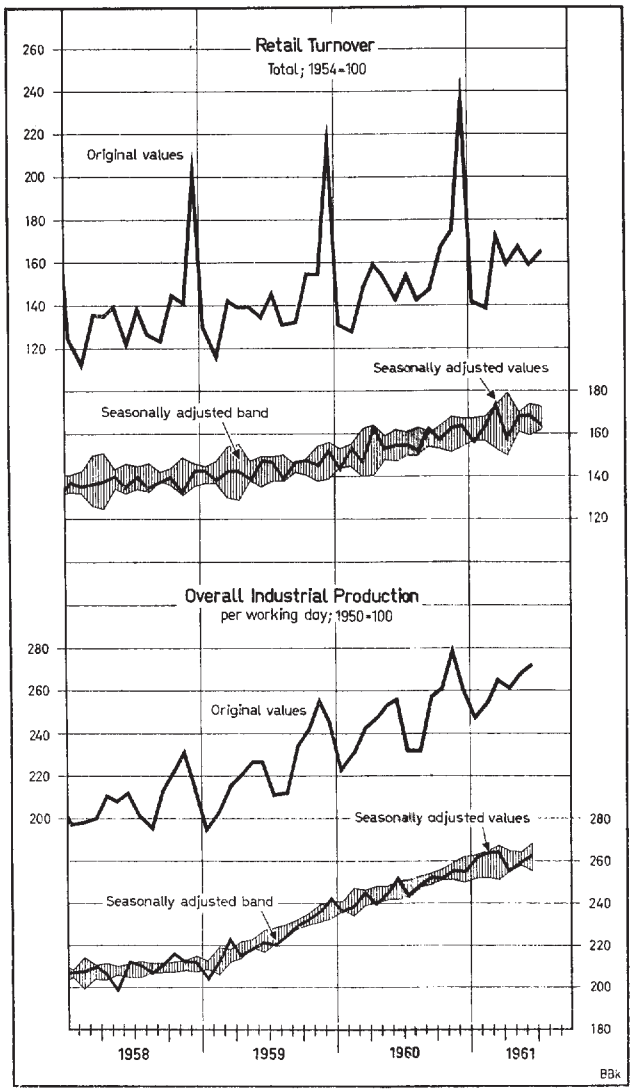
At present forty statistical series or part-series are seasonally adjusted by the Bank's Department of Research and Statistics. They mostly relate to retail turnover, to the index of industrial production and to the orders reaching industry, as well as to foreign-trade and unemployment figures. Consideration of the original values for these series shows their seasonal behaviour to be highly varied. Whereas in the case of retail turnover and the production of all industries (Graph 1) a marked seasonal movement can be discerned over a number of years, any such regularity is for example less clear in the case of incoming orders and of import values (Graph 2).

The disturbance components are relatively small in the first two cases, and relatively great in the other two.

Let us look first at the two series with a clearly discernible seasonal behaviour. In the case of total retail turnover the sharp monthly fluctuations are appreciably reduced through seasonal adjustment. Thus

*Graph 1*

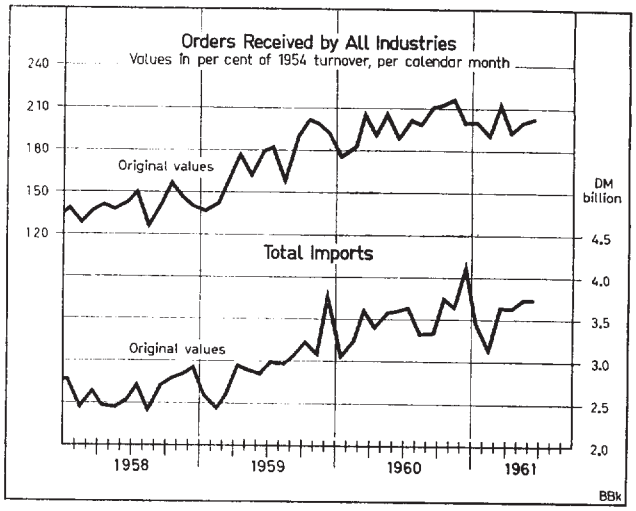
Clearly Discernible Seasonal Movements




for example the pronounced December peaks of the original series can no longer be seen in the seasonally adjusted values. For December 1960 about 95 per cent of the difference between original value and trend was eliminated. The seasonal behaviour in December is so exact that on application of the regression line formed from the values for the years 1954 to 1959 the mean-square deviation of the seasonally adjusted values from the trend amounts in that month to only 3.2 points, although for instance the seasonal component of the 1960 December value works out at about 82 points.

*Graph 2*

Seasonal Movements Not Clearly Discernible




Application of the procedure to the monthly values for the index of industrial production leads to a similar result. Here too the spread of the series' movement in individual months is appreciably narrowed. The pronounced peak of production occurring each November is no longer discernible in the seasonally adjusted values.

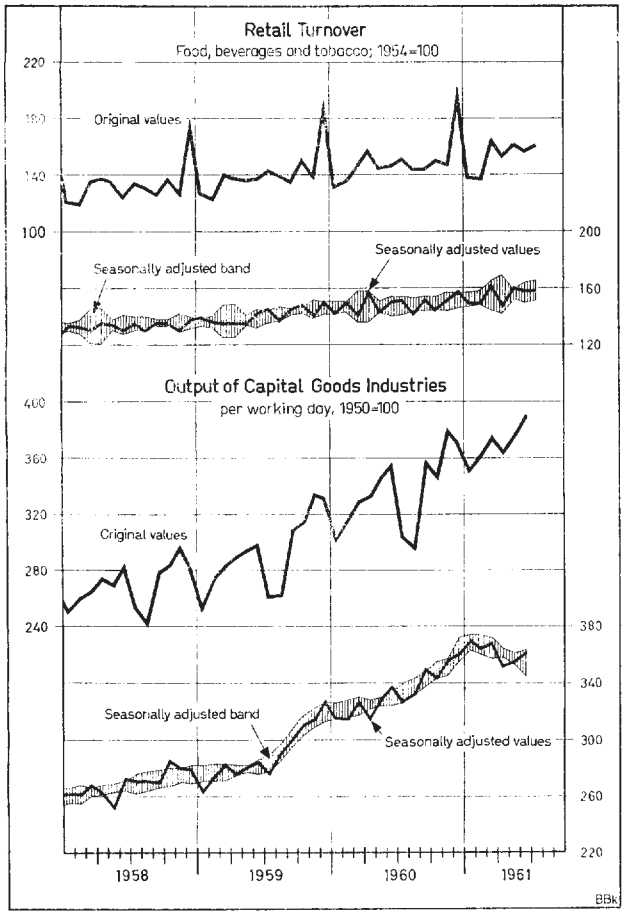
On an overall view one obtains through seasonal adjustment of the monthly values for retail turnover and for the industrial production index, if not a series free from all disturbance, at all events values in which the general direction of movement emerges appreciably better than in the original series. This applies not only to the eliminating of seasonal fluctuations from the overall series, but also to the adjustment of their most important main groups or sub-groups, such for instance as retail turnover in foodstuffs, beverages and tobacco, and the production of the capital goods industries (Graph 3).

In order to check the results the seasonally adjusted values for the main groups are combined, with weighting according to their share in the total original series, to form a seasonally adjusted overall series. The seasonally adjusted values (ascertained independently of this) for the total original series differ in most cases only slightly from the values for this series.

By way of contrast to this, for instance the seasonal adjustment of the monthly original figures for orders reaching all industries and for import values (Graph 2) yields no satisfactory results. With relatively great disturbance components in the individual months the original series are not materially smoothed by seasonal adjustment. A seasonal movement conforming to the

*Graph 3*

Seasonal Adjustment in Sub-groups of Overall Series

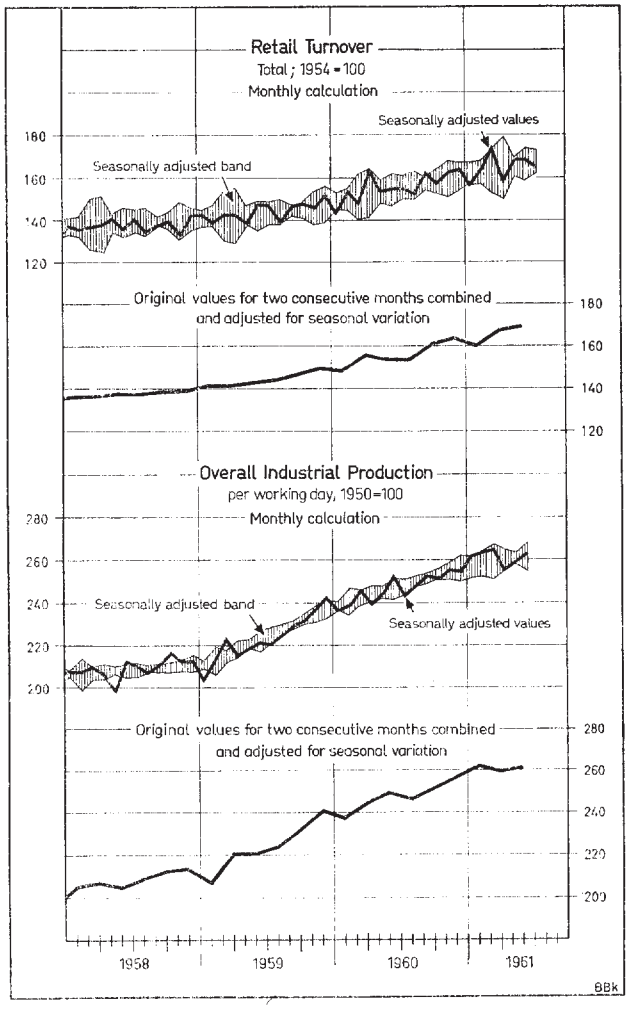



month-to-month variation can accordingly not be established.

The attempt at seasonal adjustment of most of the series did however lead to much better results when the starting point was no longer the individual monthly figures, but when, instead, two monthly values were in each case combined in a particular manner before adjustment. That is mainly because certain factors causing the seasonal movement do not operate at precisely the same time in every year. An example is afforded by retail turnover. A material influence on the size of turnover in retail trade is for instance exerted by the winter clearance sales, which begin on the last Monday in each January and last for a fortnight. Consequently in individual years a differing number of clearance-sale days falls in January and in February. A likewise perceptible effect on retail turnover is produced by Easter, which falls either in March or in April. The same applies to Whitsun, which is celebrated in May or June. In July and August the same is true of the summer clearance sales as was said about the winter clearance sales in regard to the months of January and February. While for September and October no special interdependence was found, November and December are linked by the Christmas trade, of which in every year a portion differing in size relates to the individual months. Through the fact that the seasonal adjustment was carried out for the *average* of the original values for the first and second, third and fourth months and so on in each year, the result as regards total retail turnovers was a fairly well balanced and only slightly disturbed line of seasonally adjusted values (Graph 4). The mean-square deviation of the seasonally adjusted values from the trend, which deviation after the monthly adjustment still showed values of 2 to 6 points in individual months, dropped to as low as 0.5 for the two-month period January/February. The highest value was 2.6 points for the average of the months September and October.

*Graph 4*

Seasonal Adjustment
of Combined Values for Two Consecutive Months
as Compared with Monthly Calculation
for Series with Clearly Discernible Seasonal Movements

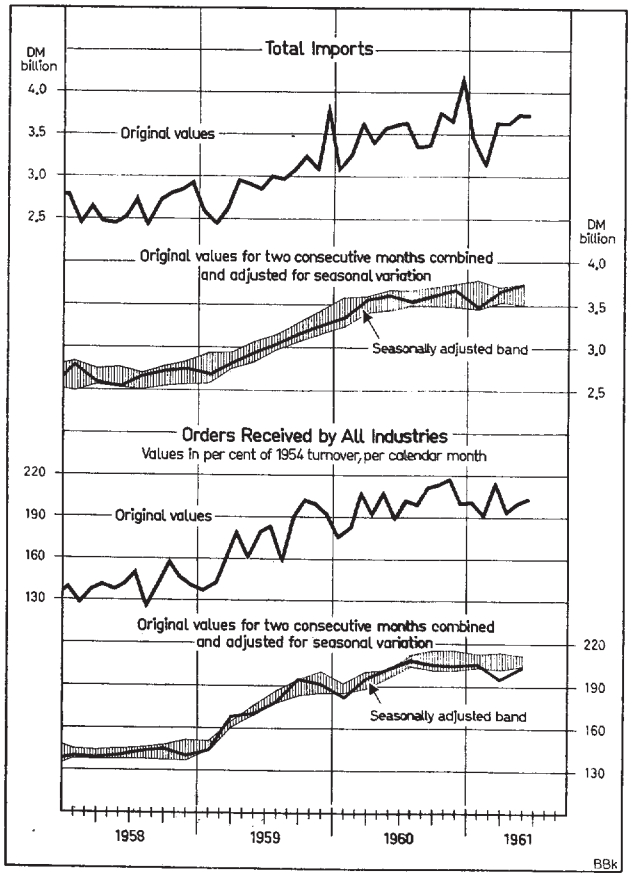


In the case of the industrial production index a similarly effected seasonal adjustment also produced a further smoothing of the series, so that here again the tendency of the general movement became still more clearly apparent. But in the case of the production index, unlike the retail turnover series, it is not immediately obvious which of the monthly values in the year should be paired together. For this reason both the sequence January/February, March/April, etc., and the sequence December/January, February/March were calculated, and then the first sequence was given preference, because it showed the smaller mean-square deviations.

Whereas combining two monthly values of the original series at a time produced in these series an improvement of the intrinsically not bad results of monthly seasonal adjustment, it was only through the same measure that effective eliminating of seasonal fluctuations from the series for orders received by industry and for the foreign trade figures was rendered possible at all.

*Graph 5*

Seasonal Adjustment
of Combined Values for Two Consecutive Months
for Series with Seasonal Movements Not Clearly Discernible




Here again the two possible pairings of the year's successive monthly values were worked out, and the series with the smaller residues was used; it was once more the sequence January/February and so forth (Graph 5).

Although in the case of retail turnover and of industrial production the seasonal behaviour is much more exact than in that of incoming orders and of imports, there is nevertheless apparent in the two last-mentioned series a certain seasonal movement which after the combining of two monthly values of the original series at a time can be satisfactorily eliminated with the aid of the regression method.

### 2. Treatment of Difficult Series

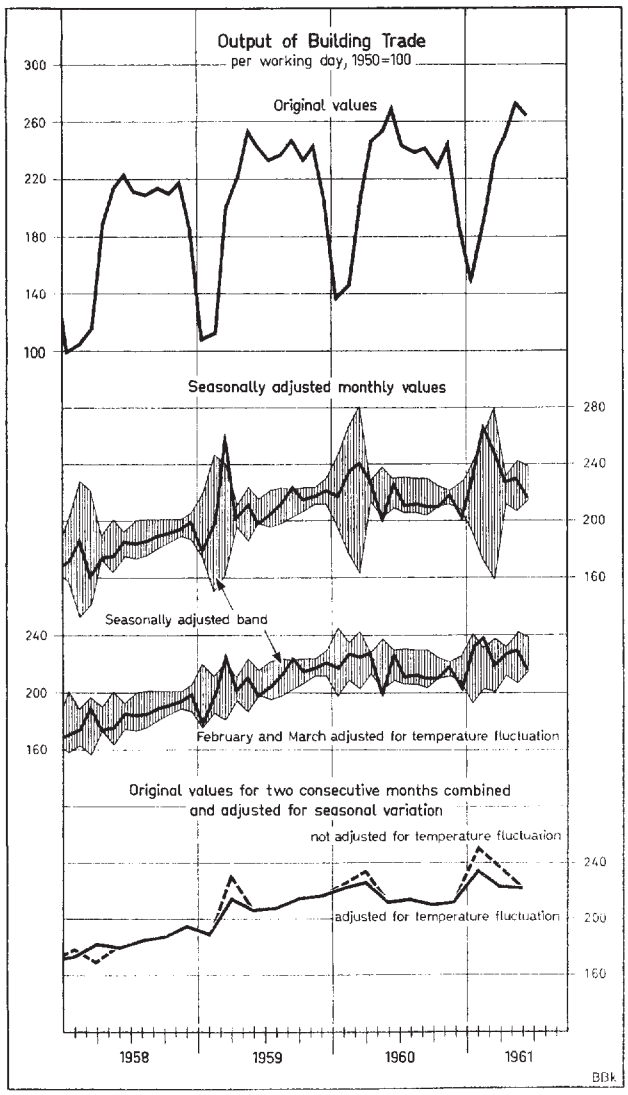
There are however series which do show a typical seasonal pattern, but in connection with which the regression procedure in the form hitherto described *nevertheless* does not yield good results. Among the series seasonally adjusted in the Bank this applies to the output of the building industry and to the unemployment figure. The peaks still present after monthly adjustment were as usual largely removed by combined adjustment of two monthly values, except for the peaks in February and March, where the seasonal component is virtually not dependent on the trend. In their case the original procedure, in which the dependence of the seasonal component on the trend value is exclusively assumed, proves a failure. On the seasonal adjustment of both monthly and two-monthly values there proved to be relatively great disturbance components, and consequently very wide dispersion bands.

The following procedure was therefore adopted:

Since it could be assumed that the differing seasonal deviations in individual years were due to the differing degree of winter cold, the residual components for the months of February and March were first correlated with the deviations shown by the temperature in those months from a weighted monthly average for a number of years. When it was thereupon found that the correlation between residual component and temperature deviation in those months was very close, a regression equation between the two was calculated, and with its help the differing influence of temperature was eliminated. Thereby it proved possible to obtain materially better seasonally adjusted values for January and February, which is indeed reflected in the narrower dispersion band (Graph 6).

*Graph 6*

Eliminating the Varying Influence of Temperature in Winter




### 3. Allowance for Seasonal Changes

As already mentioned in the preliminary note, the seasonal adjustment procedure is applied for moving periods. The parameters of the regression equations are calculated afresh each year, the period taken as basis (seven years) being moved one year forward each time. This procedure makes it possible to approximate quite closely not only a linear relation but also certain non-linear connections—for instance slight bends of a hypothetical regression line only upwards or only downwards — and to do so in particular when the trend changes only in one direction. The experience gained with moving-period application of the regression analysis is however as yet very small. The available time-series, which since the removal of governmental controls again show a true seasonal movement, seldom cover more than nine years. The parameters of the regression equations were in general calculated afresh only twice, namely — after the 1952/58 initial period — for the periods 1953/59 and 1954/60. Major changes in the direction of the trend occur during this period only in the case of a few series, for instance in the production indices for coal mining (1958/59) and for the textile industry (1957/58).

Statements as to the series for which the moving-period procedure is appropriate, and those for which it is not, cannot be made on the basis of the relatively short period; they must be reserved for a later date. One point did however emerge with great clarity, namely that it is particularly important to apply an accurate check to extremely situated points in the regression diagram, and in certain circumstances appropriately to correct them before recalculation of the regression line, so that in this way a regression line having a good course may be obtained.

If the seasonal fluctuations change abruptly, then it is desirable to shorten the calculation periods. This may be illustrated by the example of retail sales of clothing, underwear, house linen and footwear. For these even the series seasonally adjusted after the pairing of two months' values at a time continued to show, especially in the two-monthly periods September/October and November/December in the years 1957 to 1960, uniform fluctuations which pointed to the conclusion that the seasonal influences had altered (Graph 7).

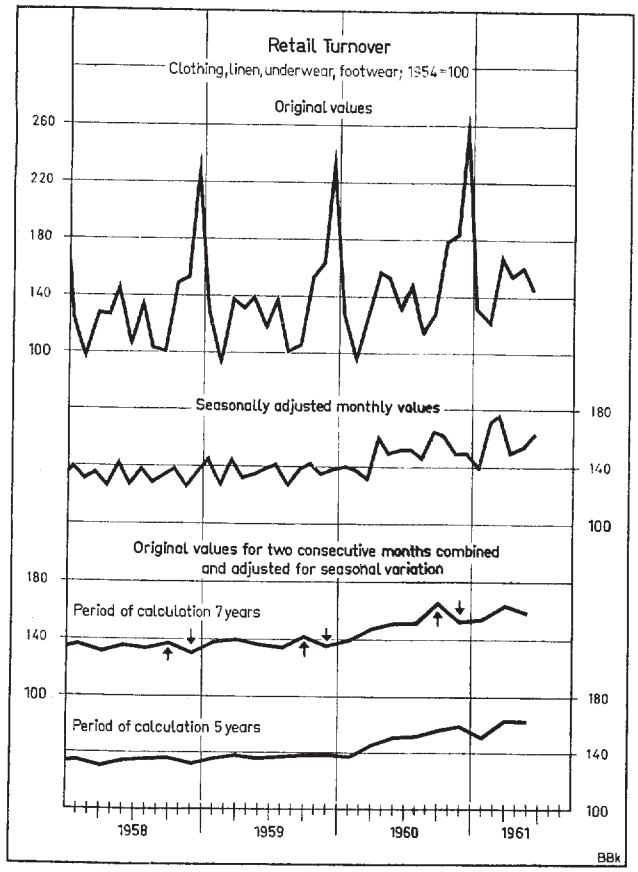
So as to obtain a new regression line better adapted to the changed seasonal behaviour, therefore, the 1954/60 period last used for computing the regression equation was replaced by the period, reduced by two years, from 1956 to 1960. With the new regression equations for the individual two-month periods the series for 1957 — the year when the seasonal change began — and subsequent years was adjusted afresh. This yielded a series in which the wide deviations previously observed have almost completely disappeared. In order to make the greatest possible allowance for such changes in the seasonal pattern it is desirable always to calculate back, with the new regression equation, as far as the time when the pattern altered.

### 4. Interpretation of the Last Value Computed

Since the latest value in a time-series is not generally included in the regression equation, its parameters can in all cases be computed with securely established trend values. For seasonal adjustment, however, an extrapolated trend has to be used. If it later turns out that the extrapolated trend was obviously wrong, and if the

*Graph 7*

Effects of Shortening the Period of Calculation




trend value for the month in question is corrected, then the seasonally adjusted value also changes; and the greater the trend alteration is, and the more the value of the multiplicative parameter in the individual regression equation differs from 1, the more it will change.

It is however found that on the whole this value lies in the neighbourhood of 1, so that even after major trend corrections the new seasonally adjusted value differs in general only immaterially from the originally calculated value. To illustrate the orders of magnitude the following example may be given. The multiplicative parameter of the regression equation, with which the May 1961 value of the production index for all industries was adjusted, amounted to 1.05. On correcting by for instance + 10 per cent (that is 26 index-points) the trend value used in seasonal adjustment for May 1961 the seasonally adjusted value would fall by 0.5 per cent (that is 1.4 index-points). In general it is found for this case that on a change of the trend value by a specified percentage x the seasonally adjusted value will alter (in the opposite direction) by 5 per cent of x.

There are however some extreme cases in which the multiplicative parameter of the regression equation is relatively great; here the effect of a trend correction is definitely perceptible, so that in such cases the seasonally adjusted value is no longer very reliable. The multiplicative parameter of the regression equation, with which for instance the March 1961 value of the building output index was seasonally adjusted, amounted to 1.51. On a trend translation of say + 10 per cent (that is 22 points) the seasonally adjusted value would here fall by 4.5 per cent (that is 11.5 points); or, in general terms, on a change of the trend value by x per cent the seasonally adjusted value would in this case alter (in the opposite direction) by 45 per cent of x. A last value so uncertainly adjusted permits only highly uncertain statements regarding the direction of movement in a series. Such cases are however extremely rare in the statistical material handled by the Bank. Nearly 70 per cent of the coefficients in the regression equations lie between 0.9 and 1.1, and roughly 90 per cent of them lie between 0.8 and 1.2. In addition it hardly happens that for several months in succession the parameters differ very widely from 1; in most cases the seasonally adjusted value for the next following month is already relatively reliable.