REMOVING THE RESIDUAL IN STANDARDIZATION PROCEDURE

Ravindra H. Dholakia*

In regional analysis, one of the most important problems is to distribute observed change in a (dependent) variable, such as employment, output or productivity, among various component factors or (independent) variables, such as industry mix, output per worker, etc. This change could be at a point of time between the regional unit and the national unit, or over a period of time. The technique most commonly used is the standardization procedure. The very popular Shift-Share analysis can be viewed only as an application of more general standardization procedure. This is directly or indirectly recognized by H. J. Brown (1969), Norcliffe (1977), Stevens and Moore (1980), Gorden et al. (1980), Ireland and Moomaw (1981), etc. As pointed out by Denison (1957), there are two ways of finding out the contribution of different factors to the observed change in the dependent variable. We may call them the partial contribution approach and the total contribution approach, respectively.

Since observed changes in the factors are never insignificantly small, interaction among these changes becomes significant and appears as residual in both these approaches. The existence of the residual makes the estimate of contribution by a given factor non-unique.

To illustrate, let us consider the following simple hypothetical example out of a wide range of the possible applications of this method:

<table>
<thead>
<tr>
<th>WPR</th>
<th>p_a</th>
<th>p_b</th>
<th>Str. Resid.</th>
<th>(PCP_A - PCP_N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>State A</td>
<td>0.5</td>
<td>400</td>
<td>250</td>
<td>0.5</td>
</tr>
<tr>
<td>Nation</td>
<td>0.6</td>
<td>250</td>
<td>0.6</td>
<td>750</td>
</tr>
</tbody>
</table>

Where WPR is overall worker population ratio; \(1_a\) and \(1_b\) are proportions of working force in sectors a and b; \(p_a\) and \(p_b\) are the productivity per worker (in dollars) in sectors a and b; and PCP is the per capita production (in dollars). The problem is finding out the contributions of WPR, \(p_a\), \(p_b\) and industrial structure in the observed deviation of state A’s PCP from that of the nation. Taking the nation as the base, we may calculate the contributions of different factors (in dollars) by the above-mentioned two approaches. These contributions (in dollars) turn out to be as follows:

<table>
<thead>
<tr>
<th>WPR</th>
<th>p_a</th>
<th>p_b</th>
<th>Ind. Resid.</th>
<th>(PCP_A - PCP_N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partial contribution</td>
<td>+45</td>
<td>+45</td>
<td>-100</td>
<td>+25</td>
</tr>
<tr>
<td>Total contribution</td>
<td>+32.5</td>
<td>+45</td>
<td>-150</td>
<td>-9</td>
</tr>
</tbody>
</table>

It can be readily seen from these figures that, by and large, the contribution of a factor depends on the approach followed. Sometimes, as is the case with the industrial structure in our example, even the sign of the contribution differs between the two approaches. To overcome this type of non-uniqueness of the contributions of the given factors, several economists have suggested methods to get the contributions of the factors such that the residual is removed. Denison (1957), for example, advocates a simple average of the two types of contributions for each factor. Thirlwall (1969) advises distributing the residual equally among the factors. Deane (1953) resolves the problem by studying the ratio rather than the difference and by applying the technique of index numbers. Farooq (1973) also resolves the problem in much the same way, by taking the geometric mean of the two indices. In principle, the last two methods are equivalent to the one suggested by Denison (1957). On the other hand, Chenery (1960), and Lewis and Soligo (1965), while measuring the contribution of import substitution, and David et al. (1971), while examining the regional growth differential, use a curious mixture of the two approaches to get rid of the residual. They distribute the observed change in the dependent variable by finding partial contribution for some factors and total contribution for others. This is obviously an inconsistent procedure with an element of arbitrariness.

We calculate the contributions of different factors in our example according to the “solutions” offered by Thirlwall (1969) and Denison (1957):

*Associate Professor, Department of Economics, M.S. University of Baroda, India. I am grateful to the anonymous referees of this journal for valuable suggestions on my exposition.
Removing the Residual in Standardization Procedure

Thirlwall’s solution (P) 33.75 33.75 -111.25 13.75 0 -30
Thirlwall’s solution (T) 45.37 57.875 -137.125 3.875 0 -30
Denison’s solution 38.75 45 -125 8 +3.25 -30

Note: Thirlwall’s solutions (P) and (T) are derived respectively from the partial and total contribution approaches.

It can be readily seen that the “solution” offered by Thirlwall (1969) is too arbitrary and sometimes inconsistent. It is arbitrary because the ultimate contribution of the same factor differs depending on the approach followed to arrive at it. Sometimes it becomes inconsistent because in the case when the two approaches give exactly the same contribution for a factor (e.g., $P_a$ in our case), there is no reason why the average contribution of the same factor should be different (which it would be if Thirlwall’s method is followed). Denison’s suggestion (1957) may also not always solve the problem because, even when a simple average of the two types of contributions is taken for each factor, some residual may still appear. The only case where the “solutions” by Denison (1957) and Thirlwall (1969) coincide and give satisfactory results is that which involves only two factors (or independent variables). Thus, their method is specific to the case of two factors only. For general applicability, therefore, we require a consistent and less arbitrary new method to remove the residual.

A. J. Brown (1973) enunciates a theorem that the residuals in the two approaches are equal in magnitude and opposite in sign. If this theorem always holds, there cannot be any failure with Denison’s method. It is important to note, however, that this theorem is specific to the case of only two factors. With more than two factors, the magnitudes of the residuals need not be equal, nor need their signs be opposite. This can be shown as below:

(1) If $Y = f(x_1, x_2, x_3, \ldots, x_n)$
then, $Y = P_1 + P_2 + \ldots + P_n + R$

$P_i$ is the partial contribution of $i$th factor, and $R$ is the residual in the partial contribution approach.

Considering that $R$ results from the interactions of finite changes in the variables $x_1, x_2, \ldots, x_n$, $R$ can be further decomposed into interactions of changes among different combinations of variables:

Using equations (1) to (4), we find that

\[ T_1 + T_2 + \ldots + T_n = \Delta Y + R + \sum_{i=1}^{n-2} \sum_{j=2}^{n-1} R_{ijk} + \ldots + \sum_{k=3}^{n-1} \sum_{j=2}^{n-2} R_{ijk} + \sum_{i=1}^{n-1} R_{i2} + \ldots + (n-2) R_{12} + \ldots + (n-2) R_{12} + \ldots + n (\text{with}\ i\neq j\neq k \neq l \neq \ldots). \]

\[ \Delta Y = T_1 + T_2 + T_3 + \ldots + T_n + Q \]

Where $Q$ is the residual in the total contribution approach and $T_i$ is the total contribution of $i$th factor which is defined as

\[ T_i = P_i + \sum_{j=1}^{n-1} R_{ij} + \sum_{j=2}^{n-1} R_{ijk} + \ldots + R_{i23} + \ldots + n (\text{with}\ i \neq j \neq k \neq l \neq \ldots). \]

In other words, $T_i$ includes $P_i$ besides all the interaction terms containing $i$ from (2) above. This is because in the total contribution approach, we make all other variables change except $x_i$. As a result, interaction of changes in all other variables with changes in $x_i$ gets included in $T_i$.

Now using equations (1) to (4), we find that

\[ T_1 + T_2 + \ldots + T_n = \Delta Y + R + \sum_{i=1}^{n-2} \sum_{j=2}^{n-1} R_{ijk} + \sum_{k=3}^{n-1} \sum_{j=2}^{n-2} R_{ijk} + \ldots + \sum_{i=1}^{n-1} R_{i2} + \ldots + (n-2) R_{12} + \ldots + n (\text{with}\ i \neq j \neq k \neq l \neq \ldots). \]

Using (3) above, we get

\[ n - 2 n - 1 \]

\[ n - 3 n - 2 n - 1 n \]

\[ \sum_{i=1}^{n-1} R_{i2} + \sum_{j=2}^{n-2} R_{ijk} + \sum_{k=3}^{n-1} R_{ijk} + \ldots + \sum_{i=1}^{n-1} \sum_{j=2}^{n-2} \sum_{k=3}^{n-1} R_{ijk} + \ldots + (n-2) R_{12} + \ldots + n \]

\[ n - 2 n - 1 n \]

\[ n - 3 n - 2 n - 1 n \]

\[ \sum_{i=1}^{n-1} \sum_{j=2}^{n-2} \sum_{k=3}^{n-1} R_{ijk} + \ldots + (n-2) R_{12} + \ldots + n \]

\[ (\text{with}\ i \neq j \neq k \neq l \neq \ldots). \]
From equation (5), it becomes clear that \( R \) and \( Q \) need not be of equal magnitude nor have opposite signs. If, however, there are only two factors, the equation (5) yields \(-Q = R\) and the Brown theorem holds. It should also be noted that with more than two factors we may get \(-Q = R\) as a special case and that Denison's method (1957) may work. But in general, Denison's method will not work. In order to have a general solution, the following method is suggested.

If we want the average value of the residual to be zero, we can find the corresponding weights for the two residuals obtained from the partial and total contribution approaches:

Thus, \( KQ + (1-K)R = 0 \) where \( K \) is the weight attached to \( Q \).

\[
K = \frac{R}{R-Q} \quad \text{and} \quad 1-K = \frac{-Q}{R-Q}
\]

If we assume that these weights also apply in the case of each factor, we can find out the average contributions of different factors such that they fully exhaust the observed change in \( Y \), or in other words, the residual turns out to be zero.

The following are the average contributions of the factors (in dollars) according to our solution in the above example:

<table>
<thead>
<tr>
<th>WPR</th>
<th>( p_a )</th>
<th>( p_b )</th>
<th>Ind. Resid. (PCP_A)</th>
<th>Struc. (PCP_N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our</td>
<td>+39.17</td>
<td>+45</td>
<td>-123.32</td>
<td>9.15</td>
</tr>
</tbody>
</table>
| Total contribution of a factor, on the other hand, is obtained by assuming that all other factors are allowed to change while only that particular factor remains at the same level.

Since Shift-Share analysis is a special case of the general method considered here, it should be recognized that the contributions of factors obtained from the Shift-Share analysis will not be unique unless the residual is removed.

To be more precise, \( p = \frac{\partial Y}{\partial x_i} \Delta x_i \), where \( \partial Y/\partial x_i \) is obtained from the function \( Y = f(x_1, x_2, \ldots, x_n) \).

REFERENCES


FOOTNOTES

1Partial contribution of a factor is obtained by assuming that all other factors remain at the same level while only that factor is allowed to change.