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[1] and the total contribution approach,[2] 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.[3]

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

	WPR	p_a	1_a	p_b	1_b	PCP
State A	0.6	400	0.5	250	0.5	195
Nation	0.5	250	0.6	750	0.4	225

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:

	WPR	p_a	p_b	Ind. Stru.	Resid-ual	$(PCP_A - PCP_N)$
Partial contri-bution	+45	+45	−100	+25	−45	−30
Total contri-bution	+32.5	+45	−150	− 9	+51.5	−30

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):

	WPR	p_a	p_b	Ind. Stru.	Resid-ual	$(PCP_A - PCP_N)$
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 $Y = f(x_1, x_2, x_3 \ldots \ldots x_n)$

then, $Y = P_1 + P_2 + \ldots \ldots p_n + R$

P_i is the partial contribution of ith factor,[4] and R is the residual in the partial contribution approach.

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

$$(2)\ R = \sum_{i=1}^{n-1} \sum_{j=2}^{n} R_{ij} + \sum_{i=1}^{n-2} \sum_{j=2}^{n-1} \sum_{k=3}^{n} R_{ijk} + \sum_{i=1}^{n-3} \sum_{j=2}^{n-2} \sum_{k=3}^{n-1} \sum_{l=4}^{n} R_{ijkl} + \ldots + R_{1 \cdot 2 \cdot 3 \cdot \ldots \ldots n} \text{ (with } i \neq j \neq k \neq l \neq \ldots)$$

Here on the R.H.S., the first term denotes the summation of interactions or changes among all possible combinations (nC_2) of only two variables; the second term denotes summation of interactions of changes among all possible combinations (nC_3) of three variables; and so on. The last term on the R.H.S. of (2) is the extent of interaction of changes among all n variables (since $nCn = 1$).

In the toal contribution approach, we get

$$(3)\ \triangle Y = T_1 + T_2 + T_3 + \ldots \ldots + T_n + Q$$

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

$$(4)\ T_i = P_i + \sum_{j=1}^{n} R_{ij} + \sum_{j=1}^{n-1} \sum_{k=2}^{n} R_{ijk} + \ldots + R_{1 \cdot 2 \cdot 3 \cdot \ldots n} \text{ (with } i \neq j \neq k \neq \ldots)$$

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

$$T_1 + T_2 + \ldots \ldots + T_n = \triangle Y + R + \sum_{i=1}^{n-2} \sum_{j=2}^{n-1} \sum_{k=3}^{n} R_{ijk} + 2 \sum_{i=1}^{n-3} \sum_{j=2}^{n-2} \sum_{k=3}^{n-1} \sum_{l=4}^{n} R_{ijkl} + \ldots$$

$$\ldots + (n-2) R_{1 \cdot 2 \cdot 3 \ldots n} \text{ (with } i \neq j \neq k \neq l \neq \ldots).$$

Using (3) above, we get

$$(5)\ -Q = R + \sum_{i=1}^{n-2} \sum_{j=2}^{n-1} \sum_{k=3}^{n} R_{ijk} + 2 \sum_{i=1}^{n-3} \sum_{j=2}^{n-2} \sum_{k=3}^{n-1} \sum_{l=4}^{n} R_{ijkl} + \ldots + (n-2) R_{1 \cdot 2 \cdot 3 \ldots n}$$

(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, $KQ + (1-K) R = O$ where K is the weight attached to Q.

$$\therefore K = \frac{R}{R-Q} \quad \text{and} \quad 1-K = \frac{-Q}{R-Q}$$

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:

	WPR	p_a	p_b	Ind. Stru.	Resid- ual	(PCP$_A$ $-$PCP$_N$)
Our Solution	+39.17	+45	−123.32	+9.15	0	−30

It is evident that our method not only yields unique contributions of the factors but also avoids the type of inconsistency which can arise in Thirlwall's method. This is because not only does the residual completely vanish in our method, but it also gives the same average contribution of a factor when the partial and the total contributions of the factor are the same (as is the case with p_a in our example).

FOOTNOTES

[1] Partial 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.

[2] 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.

[3] 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.

[4] To be more precise, $P_i = (\partial Y / \partial x_i) \triangle x_i$, where $\partial Y / \partial x_i$ is obtained from the function $Y = f(x_1, x_2, \ldots x_n)$.

REFERENCES

Brown, A. J. "Regional Economics with Special Reference to the United Kingdom." The Royal Economic Society: *Surveys of Applied Economics*, Vol. I, Macmillan, (1973).

Brown, H. J. "Shift-Share Projections of Regional Economic Growth: An Empirical Test." *Journal of Regional Science*, Vol. 9 (1969), 1–18.

Chenery, H. B. "Patterns of Industrial Growth." *American Economic Review*, Vol. L, No. 4, September (1960).

Davis, H. W., R. T. Newsom and D. E. O'Neill. "Rate-Weight Analysis: A Suggested Technique for Examining Regional-National Growth Rate Differential." *Annals of Regional Science*, Vol. V, No. 2, December (1971).

Deane, Phyllis. "Regional Variations in U.K. Incomes from Employment." *Journal of the Royal Statistical Society* (A), (1953), 123–135.

Denison, E.F. "Comment." N.B.E.R.: *Regional Income-Studies in Income and Wealth*, Vol. 21, (1957).

Farooq, G. M. "Economic Growth and Changes in the Industrial Structure of Income and Labour Force in Pakistan." *Economic Development and Cultural Change*, Vol. 21, No. 2, January (1973).

Gordon, J. R., W. B. Hackett and D. Mulkey. "Using the Shift-Share Technique in Economies with Widely Varying Sectoral Growth Rates: Observations and a Suggested Model Modification." *Review of Regional Studies*, Vol. 10, No. 1, Spring (1980).

Ireland, T. C. and R. L. Moomaw. "The Competitive Effect in Shift-Share Analysis: A Will of the Wisp?" *Review of Regional Studies*, Vol. 11, No. 1, Spring (1981).

Lewis, S. R. and R. Soligo. "Growth and Structural Change in Pakistan Manufacturing Industry 1954–64." *Pakistan Development Review*, Vol. V, No. 1, Spring (1965).

Norcliffe, G. B. "A Components Model for Disaggregating Regional Productivity Variations." *Regional Studies*, Vol. 11, No. 6 (1977).

Stevens, B. H. and C. L. Moore. "A Critical Review of the Literature on Shift-Share as a Forecasting Technique." *Journal of Regional Science*, Vol. 20, No. 4, (1980).

Thirlwall, A. P. "Weighting System and Regional Analysis; A Reply to Mr. Cunningham." *Oxford Economic Papers*, Vol. 21, March (1969).