NOTES AND COMMENTS

MULTILEVEL DECOMPOSITION OF THEIL'S INDEX OF INEQUALITY

BY FRANK A. COWELL
London School of Economics

In a recent edition of this Review, Adelman and Levy (1984) extend the work of Fishlow (1972) and argue that the multilevel decomposition of the Theil index is misleading.

The purpose of this note is to show that if one is careful about specifying one's terms, multilevel decomposition of the Theil index is actually quite straightforward. Indeed neat decomposability is one of the supremely attractive qualities of this index (see Cowell, 1980), in sharp contrast to the Gini coefficient which is only decomposable if the constituent subgroups can be strictly ordered by income. I show the elements of what is required in formal, but simple terms, and then illustrate this using Adelman and Levy's data. I also show that they are wrong on one very important point—their assertion that overall inequality depends on the way in which one does the decomposition.

As far as possible I shall make use of Fishlow's notation, but with some modifications for the sake of clarity. Let $\mathscr{I}$, $\mathscr{J}$, $\mathscr{K}$ be three different partitions of a given population. Each such partition consists of two or more index sets. For example $\mathscr{K}$ might be the partition by educational background: in the Adelman-Levy case this consists of four sets of individuals, each categorized by a particular education level. We may also construct finer partitions by combining two out of the above three. For example $\mathscr{J}$ might denote birth place (two categories in the Adelman-Levy example): in which case $\mathscr{J}\mathscr{K}$ refers to a partition by birth-place *and* by education (containing, of course, eight possible categories in Adelman-Levy). Naturally, however fine the partition, every individual in the population belongs to one and only one constituent set of the partition.

These ideas are illustrated in Figure 1, a simple Venn diagram. Let the big circle represent the entire population; let ethnic groups be partitioned off by vertical strips, birth-place groups be partitioned off by horizontal strips and educational groups by eccentric circles. Then (using the Adelman Levy example) we have

$$\mathscr{I} = \{\alpha \cup \beta \cup \gamma \cup \delta, \varepsilon \cup \zeta \cup \eta \cup \theta \cup \kappa \cup \lambda \cup \mu\}$$

$$\mathscr{J} = \{\alpha \cup \kappa \cup \lambda \cup \mu, \beta \cup \gamma \cup \delta \cup \varepsilon \cup \zeta \cup \eta \cup \theta\}$$

$$\mathscr{K} = \{\alpha \cup \beta \cup \varepsilon \cup \kappa, \gamma \cup \zeta \cup \lambda, \delta \cup \eta \cup \mu, \theta\}$$

$$\mathscr{I}\mathscr{J} = \{\alpha, \beta \cup \gamma \cup \delta, \varepsilon \cup \zeta \cup \eta \cup \theta, \kappa \cup \lambda \cup \mu\}$$

$$\mathscr{I}\mathscr{K} = \{\alpha \cup \beta, \gamma, \delta, \varepsilon \cup \kappa, \zeta \cup \lambda, \eta \cup \mu, \theta\}$$

$$\mathscr{J}\mathscr{K} = \{\alpha \cup \kappa, \beta \cup \varepsilon, \gamma \cup \zeta, \delta \cup \eta, \theta, \lambda, \mu\}.$$

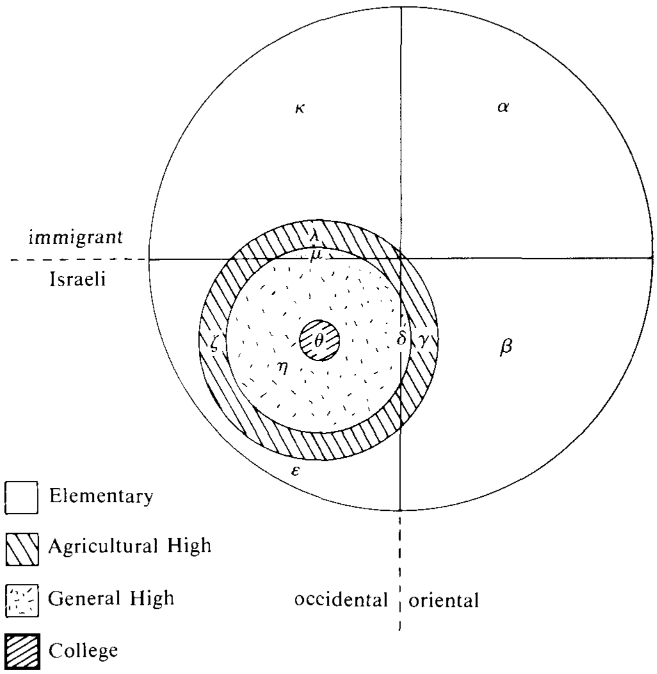

Figure 1

Notice that some of the theoretically possible member sets of the partition are in fact null: for example the "college" category is a strict subset of "Israeli-occidental."

Now let i, j, k index the sets in the basic partitions $\mathscr{I}$, $\mathscr{J}$, $\mathscr{K}$ and let X_{ijk} and Y_{ijk} be, respectively the number of individuals in and the aggregate income received by subgroup ijk (in the example these quantities represent the numbers and incomes associated with each basic non-null cell $\alpha, \ldots, \mu$. This permits one to define "shares" in an appropriate fashion:

$$x_{ijk} = X_{ijk}/\Sigma_{i'}\Sigma_{j'}\Sigma_{k'}X_{i'j'k'}$$

$$Y_{ijk} = Y_{ijk}/\Sigma_{i'}\Sigma_{j'}\Sigma_{k'}Y_{i'j'k'}$$

where the summations are taken over the entire range of the index in each case. From these elementary concepts we may also write

$$x_{ij.} = \Sigma_k x_{ijk}$$

$$x_{i..} = \Sigma_j x_{ij.}$$

$$x_{.jk} = \Sigma_i x_{ijk}$$

$$x_{..k} = \Sigma_j x_{.jk}$$

$$x_{i.k} = \Sigma_j x_{ijk}$$

$$x_{.j.} = \Sigma_i x_{ij.} = \Sigma_k x_{.jk'}$$

with corresponding definitions for the y's.

For the Theil index, write inequality in the total population as T; write inequality *within* group i as T_i or within subgroup ij as T_{ij}; and write *between* group inequality for partition $\mathcal{I}$ as $\bar{T}(\mathcal{I})$—this is the inequality that would result if every member of the population received the mean income of the set to which he belongs under partition $\mathcal{I}$. Then (see Theil, 1967) we may write:*

(1) $$T = \Sigma_i y_{i..} T_i + \bar{T}(\mathcal{I})$$

(2) $$= \Sigma_i \Sigma_j y_{ij.} T_{ij} + \bar{T}(\mathcal{I}\mathcal{J})$$

where $T_{ij} = \Sigma_k y_{ijk} \log (y_{ijk}/x_{ijk})$. Alternatively one may of course write

(3) $$T = \Sigma_k y_{..k} T_k + \bar{T}(\mathcal{K})$$

(4) $$= \Sigma_k \Sigma_j y_{.jk} T_{jk} + \bar{T}(\mathcal{J}\mathcal{K})$$

where

$$T_{jk} = \Sigma_i y_{ijk} \log (y_{ijk}/x_{ijk}).$$

It is important to note (a) that for any given population (1) to (4) all have the same value, and (b) that (1) and (2), or alternatively (3) and (4), yield all the information one needs for multilevel decomposition.

TABLE 1

POPULATION AND INCOME SHARES FOR COMPONENT SUBGROUPS

	k = Elementary	k = Agr. High	k = Gen. High	k = College
		(i = Oriental, j = Immigrant)		
x_{ijk}	0.3897	—	—	—
y_{ijk}	0.3004	—	—	—
		(i = Oriental, j = Israeli)		
x_{ijk}	0.0164	0.0164	0.0164	—
y_{ijk}	0.0156	0.0203	0.0177	—
		(i = Occidental, j = Israeli)		
x_{ijk}	0.0376	0.0329	0.1244	0.0070
y_{ijk}	0.0407	0.0463	0.1526	0.0102
		(i = Occidental, j = Immigrant)		
x_{ijk}	0.0986	0.0469	0.2136	—
y_{ijk}	0.0973	0.0602	0.2385	—

The Adelman-Levy data have been reproduced in the form of shares in Table 1. In Table 2 I present the basic computations one needs to look at the subgroup contributions to inequality, in the manner suggested by Adelman and Levy. The contributions to inequality are then presented in Table 3.

*In this notation, of course, if the data are presented in grouped form, as in the Adelman-Levy example, then T is estimated by $\bar{T}(\mathcal{I}\mathcal{J}\mathcal{K})$.

TABLE 2

VALUES OF THEIL'S INEQUALITY INDEX FOR THE WHOLE POPULATION AND BETWEEN CERTAIN COMPONENT SUBGROUPS

T	0.0209
$\bar{T}(\mathcal{I}\mathcal{J})$	0.0184
$\bar{T}(\mathcal{J}\mathcal{K})$	0.0184
$\bar{T}(\mathcal{K}\mathcal{J})$	0.0199
$\bar{T}(\mathcal{I})$	0.0149
$\bar{T}(\mathcal{J})$	0.0070
$\bar{T}(\mathcal{K})$	0.0166

Does it matter how the decomposition is carried out? Obviously if the data and the classifications chosen happened to have the property that the partitions were "orthogonal", then further simplication is possible. Orthogonality in this case implies $\bar{T}(\mathcal{I}\mathcal{J}) = \bar{T}(\mathcal{I}) + \bar{T}(\mathcal{J})$ etc., so that the second column of part (a) of Table 3 would then simply contain the entries $\bar{T}(\mathcal{I})$, $\bar{T}(\mathcal{J})$, $\bar{T}(\mathcal{K})$ and part (b) would contain $\bar{T}(\mathcal{K})$, $\bar{T}(\mathcal{J})$, $\bar{T}(\mathcal{I})$. Clearly in this case the partitions are not orthogonal, as one may check by computing the interaction terms from Table 1 (for example $\bar{T}(\mathcal{I}\mathcal{J}) - \bar{T}(\mathcal{I}) - \bar{T}(\mathcal{J}) = -0.0035$). But this really represents no more of a problem than would correlation between different variables in analysis of variance. One merely needs to be careful how one describes the entries in a table such as Table 3. For example the first entry in the right hand column of Table 3(a) gives the amount of inequality that there would be if all inequality were eliminated within ethnic groups. The third entry in Table 3(b) gives the increase in inequality that would arise if you did not eliminate inequality within nativity-schooling subgroups. These two concepts are not generally identical, even if the only possible sources of inequality are the three factors mentioned. So there should not be any confusion about the meaning of the decomposition as one might have inferred from Adelman-Levy's paper.

TABLE 3

DECOMPOSITION OF THEIL'S INEQUALITY INDEX BY (a) ETHNICITY, NATIVITY AND LEVEL OF SCHOOLING, AND (b) LEVEL OF SCHOOLING, NATIVITY AND ETHNICITY

(a)		
Inequality between ethnic groups	$\bar{T}(\mathcal{I})$	0.0149
Inequality attributable to nativity variation within ethnic groups	$\bar{T}(\mathcal{I}\mathcal{J}) - \bar{T}(\mathcal{I})$	0.0035
Inequality attributable to schooling levels within ethnicity nativity groups	$T - \bar{T}(\mathcal{I}\mathcal{J})$	0.0025
Total Inequality		0.0209
(b)		
Inequality between schooling level groups	$\bar{T}(\mathcal{K})$	0.0166
Inequality attributable to nativity variation within schooling level groups	$\bar{T}(\mathcal{J}\mathcal{K}) - \bar{T}(\mathcal{K})$	0.0017
Inequality attributable to ethnicity within schooling nativity groups	$T - \bar{T}(\mathcal{J}\mathcal{K})$	0.0025
		0.0209

Finally note that Adelman-Levy state that "The computed magnitude of total inequality is also sensitive to the order of the decomposition." It is not. As demonstrated in my Table 3, the total must be the same whichever way you choose to do your partitions.

Bibliography

Adelman, I. and Levy, A. (1984), Decomposing Theil's Index of Inequality into Between and Within Components: a note, *Review of Income and Wealth*, 30, 119-121.

Cowell, F. A. (1980), On the Structure of Additive Inequality Measures, *Review of Economic Studies*, 47, 521-531.

Fishlow, A. (1972). Brazilian Size Distribution of Income, *American Economic Review, Papers and Proceedings*, 62, 381-402.

Theil, H. (1967), *Economics and Information Theory*, North Holland, Amsterdam.