NOTES

DECOMPOSING THEIL'S INCOME OF INCOME INEQUALITY INTO BETWEEN AND WITHIN COMPONENTS: A NOTE

BY IRMA ADELMAN AND AMNON LEVY
University of California, Berkeley

In his study of the Brazilian size distribution of income, Fishlow [1972] modified Theil's index of income inequality by decomposing the general measure to quantify the socioeconomic effects leading to inequality among workers. In particular, he considered two ways of decomposing total inequality. One is into between and within components; the other is into the contribution to total inequality of variations in mean incomes among sectors, regions, etc., taking account of interaction effects. Our comment concerns the first method.

In this note, we show that decomposing Theil's index of total inequality into between and within components cannot lead to an unambiguous quantification of the causes of inequality because, in the presence of intercorrelation between the principles of decomposition, the decomposition overstates the contribution of the first cause considered and understates the contribution of subsequent effects.

The first step in computing inequality between groups is analogous to estimating a misspecified regression equation where all but one of the explanatory variables are omitted. As is well known, if the true relationship is

$$(1) \qquad y = \sum_{i=1}^{k} b_i x_i + \varepsilon \qquad k > 1, \quad Ex_i'\varepsilon = 0 \quad \forall i = 1, \ldots, k$$

and the estimated specification is

$$(2) \qquad y = cx_1 + u,$$

then the bias of the ordinary least-square estimate of $c(\hat{c})$ is

$$(3) \qquad E(\hat{c} - b_1) = \sum_{i=2}^{k} b_i \rho_{1i}^2 \operatorname{var}(x_i)$$

where ρ_{1i} is the correlation coefficient getween x_1 and x_i. (E is the expectation operator.) If $b_i \geq 0$, $\operatorname{var}(x_i) \geq 0$, $\rho_{1i}^2 \geq 0$ for all $i = 2, \ldots, k$, and there exists at least one variable $2 \leq j \leq k$ such that $b_j > 0$, $\operatorname{var}(x_j) > 0$, $\rho_{1j}^2 > 0$, then $\hat{c}$ is overstated. The higher the correlations between the included variable and the excluded ones and the larger the variances of the excluded variables, the larger the bias of $\hat{c}$.

An analogous but not identical problem arises in decompositions of the Gini coefficient in terms of the factor components of total income of the family (Fei, Ranis, and Kuo, 1978; Pyatt, Chen, and Fei, 1980). The problem with the latter is not one of missing variables but rather of correlation among the factor Ginis.

The magnitude of the bias due to this deficiency in the decomposition of the Theil inequality index can best be demonstrated by an empirical example. Our example concerns the income distribution within 426 hypothetical family farms in Israel, 1975.[1] This population is divided into 11 groups according to the level of schooling (elementary, general high school, agricultural high school, and college), the ethnicity (Oriental, Occidental), and the nativity (immigrant, Israeli born) of the farm managers. Table 1 describes the income distribution within this population. It also shows that there are high correlations between being a farmer of Oriental ethnicity and being an immigrant and finishing elementary school only; and between being a farmer of Occidental ethnicity and finishing general high school. Table 2 summarizes the computed income inequality attributable to variations in ethnicity, nativity, and schooling of the farm managers with the decomposition of Theil's inequality index performed in this order. Table 3 summarizes the computed inequality by components with the decomposition order reversed—schooling, nativity, ethnicity. The numbers in parentheses are percent of total inequality. They demonstrate our argument that in the presence

TABLE 1

SIMULATED INCOME DISTRIBUTION UNDER UNIFORM LAND ALLOCATION WITHIN 426 HYPOTHETICAL ISRAELI FAMILY FARMS IN 1975

	School level			
	Elementary school	Agricultural high school	General high school	College
	Oriental immigrants			
Farmer income	43,576			
Number of farmers	166			
Total income[a]	7,233,616			
	Israeli-born Oriental			
Farmer income	53,782	69,922	60,857	
Number of farmers	7	7	7	
Total income[a]	376,474	489,454	425,999	
	Israeli-born Occidental			
Farmer income	61,259	79,642	69,317	81,885
Number of farmers	16	14	53	3
Total income[a]	980,144	1,114,988	3,673,801	245,655
	Occidental immigrants			
Farmer income	55,775	72,513	63,111	
Number of farmers	42	20	91	
Total income[a]	2,342,550	1,450,260	5,743,101	

[a]Ethnicity-nativity-schooling group.

[1]The income distribution was generated by a simulation in which land was assumed to be equally divided among the hypothetical settlers and which considered the production process as a monotone transformation of a fixed-proportion production function in physical inputs—the transformation depending upon the farmer's managerial ability and a linear combination of the farmer's characteristics. This production function was estimated by Berck and Levy [1982] using a sample of 426 family farms in Israel, 1974–75, provided by the Israeli Farm Income Research Institute. The participants in the simulation have the same average, variance, and covariance of characteristics as do the sample settlers.

TABLE 2

DECOMPOSITION OF THEIL'S INEQUALITY INDEX BY ETHNICITY, NATIVITY, AND LEVEL OF SCHOOLING

Sources of inequality		
1.	Inequality between ethnic groups (Oriental vs. Occidental)	0.0065 (81 percent)
2.	Inequality attributable to nativity variation (Israeli born vs. immigrants) within ethnic groups	0.0001 (1 percent)
3.	Inequality attributable to schooling level variation within ethnicity nativity groups	0.0014 (18 percent)
	Total inequality	0.0080 (100 percent)

TABLE 3

DECOMPOSITION OF THEIL'S INEQUALITY INDEX BY LEVEL OF SCHOOLING, NATIVITY, AND ETHNICITY

Sources of inequality		
1.	Inequality between schooling level groups	0.0071 (66 percent)
2.	inequality attributable to nativity variation (Israeli born vs. immigrants) within schooling level groups	0.0006 (5 percent)
3.	inequality attributable to ethnicity variation (Oriental vs. Occidental) within schooling-nativity groups	0.0031 (29 percent)
	total inequality	0.0108 (100 percent)

of high correlations between the principles of decomposition and of positive association between the variances of the variables represented in the successive principles of decomposition and total inequality, decomposing Theil's inequality index into between and within components can be quite misleading. Furthermore, it is evident from the last line of Tables 2 and 3 that the computed magnitude of total inequality is also sensitive to the order of the decomposition.

REFERENCES

Berck, Peter and Levy, Amnon, The Costs of Equal Land Distribution; The Case of the Israeli Moshavim, unpublished manuscript, Department of Agricultural and Resource Economics, University of California, Berkeley, August, 1982.

Fishlow, Albert, Brazilian Size Distribution of Income, *American Economic Review*, LXII, 391–402, 1972.

Fei, John C. H., Ranis, Gustav, and Kuo, Shirley W. Y., Growth and the Family Distribution of Income by Factor Components, *The Quarterly Journal of Economics*, XCII, 17–53, 1978.

Pyatt, Graham, Chau-Nan Chen, and John C. H. Fei, The Distribution of Income by Factor Components, *The Quarterly Journal of Economics*, XCIV, 451–473, 1980.