Factor Components, Population Subgroups and the Computation of the Gini Index of Inequality

Author(s): Jacques Silber

Source: *The Review of Economics and Statistics*, Vol. 71, No. 1 (Feb., 1989), pp. 107-115
Published by: The MIT Press
Stable URL: http://www.jstor.org/stable/1928057
Accessed: 20/09/2013 02:51

FACTOR COMPONENTS, POPULATION SUBGROUPS AND THE COMPUTATION OF THE GINI INDEX OF INEQUALITY

Jacques Silber*

Abstract—The purpose of this study is to propose a simple technique based on matrix algebra to compute the Gini Index of Inequality, to obtain a decomposition of this index by factor components, when detailed data on the various income sources are available, to derive a breakdown of the inequality into within and between classes inequality, when the income units are grouped by income range, and to compute the contribution of the within and between groups inequality as well as that of some interaction term, when the data are classified by population subgroups.

I. Introduction

OF particular interest among the numerous works on income inequality measurement which have been published in the last fifteen years are attempts which have been made to assign inequality contributions to various components of income such as labor or property income (Fei, Ranis, and Kuo, 1978; Kakwani, 1980; Pyatt, Chen and Fei, 1980; Shorrocks, 1982; Shorrocks, 1983) or to various population subgroups (Cowell, 1980; Shorrocks, 1980; Blackorby et al., 1981; Cowell and Kuga, 1981; Das and Parikh, 1982; Shorrocks, 1984).

Among these studies those which were based on the decomposition of the Gini Index of Inequality (Concentration) included often cumbersome technical sections because the Gini Index, being a function of the income as well as of the rank of the individuals, does not lend itself easily to such a breakdown by factor components or population subgroups.

Recent efforts (Lerman and Yitzhaki, 1984 and 1985; Shalit, 1985) to derive new algorithms allowing a quicker computation of the Gini Index may have been successful in decomposing inequality by income sources (components) but they do not seem to simplify the estimation procedure when data are classified by population subgroups.

The purpose of this paper is to show that the use of a new linear operator, called the G-matrix, simplifies greatly not only the computation of the Gini Index but also its decomposition by factor components or population subgroups. The proposed approach allows also to give a clear interpretation of the interaction term which one obtains when the Gini Index is broken down by population subgroups.

The paper is organized as follows: the first section shows succinctly that the G-matrix may be derived from one of the expressions of the Gini Index of Inequality (Concentration). The next three sections indicate how to use the G-matrix to obtain respectively a decomposition of the Gini Index by factor components, income classes and population subgroups. A last section is devoted to two illustrations, one looking at the contribution of various income sources to the inequality at retirement age, the other presenting a decomposition by ethnic groups of the inequality of U.S. household incomes in 1980. The purpose of the latter section is mainly to show that the technique proposed here should be useful in any applied research on income inequality.

II. Matrix Algebra and the Gini Index

Following the work of Sen (1973) and Donaldson and Weymark (1980) it has been shown (Berrebi and Silber, 1985) that the Gini Index of Inequality I_G could be written as

$$I_G = \sum_{j=1}^{n} s_j[(n-j)/n - (j-1)/n] \qquad (1)$$

where s_j is the proportion of total income earned by the individual whose income has the j^{th} rank in the income distribution, assuming that

$$s_1 \geq s_2 \geq \cdots \geq s_j \geq \cdots \geq s_n.$$

Received for publication October 20, 1986. Revision accepted for publication April 26, 1988.

* Bar Ilan University, Israel.

This research was supported by a grant from the Ford Foundation through the Israel Foundation Trustees. An earlier version of this paper was written while the author was on leave at the University of Southern California. The author wishes to thank Z. M. Berrebi and several anonymous referees for their fruitful comments and suggestions.

Expression (1) may also be written as

$$I_G = \sum_{i=1}^{n} s_i \left[\sum_{j \geq i} (1/n) - \sum_{j \leq i} (1/n) \right]. \tag{2}$$

It has been shown recently (Berrebi and Silber, 1987) that (2) could also be written as

$$I_G = (e'Gs) \tag{3}$$

where e is a column vector of n elements which are equal to $1/n$ (e' being the corresponding row vector), s is a column vector of n elements being respectively equal to $s_1, s_2, \ldots, s_n$ and G, which could be called the G-matrix, is an n by n matrix whose elements g_{ij} are equal to -1 when $j > i$, to $+1$ when $i > j$ and to 0 when $i = j$.

Since many computer programs have subroutines for matrix multiplications, expression (3) seems to be a very simple way of computing quickly the Gini Index of Inequality.

Moreover this approach may also be used to compute an upper bound to the Gini Index when only grouped observations are available.[1]

III. Factor Components and the Gini Index

Let X_j denote individual j's total income and X_{ji} the income he receives by providing productive factor i. X_{ji} will be called the factor i component of individual j's income. The share of individual j in total income will now be denoted by $s_{j.}$ and will be written

$$s_{j.} = (X_j/X_T) \tag{4}$$

where $X_T = \sum_{j=1}^{n} X_j$, whereas the share of component i in society's total income will be denoted[2] by $s_{.i}$ and defined by

$$s_{.i} = \left(\sum_{j=1}^{n} X_{ji} \right) \bigg/ X_T. \tag{5}$$

[1] The proposed method is described in an appendix which may be obtained upon request from the author.

[2] To keep the mathematical notation as simple as possible and to avoid therefore the multiplication of characters, symbols like $s_{.i}$ and $s_{j.}$ will be used to represent either the shares of component i or of individual j, or the vector of these shares whose typical element is precisely $s_{.i}$ or $s_{j.}$. The context in which these elements will be used should allow the reader to see when the symbols represent a vector and when they refer only to elements of a vector.

Let us call s_{ji} the share of the component i of individual j in total income X_T, that is,

$$s_{ji} = X_{ji}/X_T \tag{6}$$

and define an n by $(k + 1)$ matrix S, whose first column is the vector s of the shares $s_{j.}$, whose second column is the vector $s_{.1}$ of the shares s_{j1} and whose $(i + 1)^{\text{th}}$ column is the vector $s_{.i}$ of the shares s_{ji}. The product

$$e'GS = z \tag{7}$$

is a row vector z of $(k + 1)$ elements whose first element, as indicated by (3), is the Gini Index of total income inequality. The next k elements of z may be written as

$$z_{i+1} = e'Gs_{.i} \quad (i = 1 \text{ to } k) \tag{8}$$

so that

$$\sum_{i=1}^{k} z_{i+1} = e'G(s_{.1} + \cdots + s_{.i} + s_{.k}) = e'Gs = z_1. \tag{9}$$

Let us call $v_{.i}$ the vector of the ratios $(s_{ji}/s_{.i})$ ($j = 1$ to n) so that the element v_{ji} of $v_{.i}$ is the share of individual j in the total income derived from component i. Expression (8) may then be written as

$$z_{i+1} = e'G \begin{pmatrix} (s_{1i}/s_{.i}) s_{.i} \\ \vdots \\ (s_{ni}/s_{.i}) s_{.i} \end{pmatrix}$$

$$\leftrightarrow z_{i+1} = e'Gv_{.i}(s_{.i}) = (s_{.i}) e'Gv_{.i}. \tag{10}$$

The scalar $e'Gv_{.i}$ is generally not the Gini Inequality Index G_i of the i^{th} component of income since the elements v_{ji} of $v_{.i}$ ($j = 1$ to n) may not necessarily be ranked by decreasing value as are the elements $s_{j.}$ of s. To obtain the Gini Inequality index G_i of component i one has to construct a new vector $y_{.i}$ whose elements y_{ji} ($j = 1$ to n) are the shares $(s_{ji}/s_{.i})$ previously defined, but they are not ordered according to the order of the shares $s_{j.}$ of vector s (as were the elements v_{ji} of $v_{.i}$) but according to their own rank in the vector $y_{.i}$. The Gini Index G_i for component i is therefore defined as

$$G_i = e'Gy_{.i}. \tag{11}$$

The scalar $C_i = e'Gv_{.i}$, on the other hand, has been called the Pseudo Gini of component i by Fei, Ranis and Kuo (1978) or the concentration ratio of component i by Rao (1969), Kakwani

(1980), Pyatt, Chen and Fei (1980) and Shalit (1985).

Combining (9) and (10), it can be seen that the first element z_1 of z is a weighted average of the Pseudo-Gini indices C_i (concentration rations), the weights being the shares $s_{.i}$ of the various components i in society's total income.[3] If one defines $s_{.i}C_i$ as the contribution of component i to total inequality, it can then be said that the k last elements of z represent the contribution of the various k components to total inequality.[4]

Our procedure to decompose the Gini Inequality Index among all factors that contribute to income is therefore relatively simple. It requires the construction of a matrix S whose first column refers to the shares of each individual's total income in society's total income whereas the next columns refer to the shares of each individual's income from a specific component in society's total income. All these vectors are ranked according to the (decreasing) rank of the individuals in total income. The product $e'GS$ is then a vector z whose first element is the Gini Inequality Index I_G whereas the other elements are the contributions of each component to the overall inequality I_G.

IV. Income Classes and the Gini Index

When the data on income distribution are grouped by income class (range), it may be of interest to decompose total inequality into two components, the inequality between income classes and that within income classes. Kakwani (1980) has shown that the Gini Index is equal to the sum of the "between-classes" Gini Index and of a weighted average of the "within-classes" Gini Indices. It will now be shown that such a result may also be obtained when the Gini Index is computed with the help of the G-matrix previously defined.

Let us partition this n by n G-matrix into m^2 submatrices where m is the number of income classes (e.g., deciles) used. In each income class h there are n_h individuals so that $n = \sum_{h=1}^{m} n_h$. The partitioned matrix G will therefore look like

$$G = \begin{pmatrix} G(n_1, n_1) & \dots & G(n_1, n_m) \\ \vdots & & \vdots \\ G(n_m, n_1) & \dots & G(n_m, n_m) \end{pmatrix}. \quad (12)$$

The n_h by n_h $G(n_h, n_h)$ matrices have 0 on their diagonals, (-1)'s in their upper right triangle and $(+1)$'s in their lower left triangle. The n_p by n_q $G(n_p, n_q)$ matrices, where $q > p$, have all identical elements equal to (-1) whereas the n_r by n_t $G(n_r, n_t)$ matrices, where $t < r$, have all identical elements equal to $(+1)$.

Let us similarly decompose each of the vectors e' and s into m components, respectively, called $e'(n_h)$ and $s(n_h)$, having each n_h elements. The product $e'Gs$ defined in (3) may now be written, using well-known rules on partitioned matrices, as

$$e'Gs = \sum_{p=1}^{m} \left[\sum_{q=1}^{m} e'(n_p) G(n_p, n_q) s(n_q) \right]. \quad (13)$$

Expression (13) may also be written as

$$\begin{aligned} e'Gs &= \sum_{p=1}^{m} e'(n_p) G(n_p, n_p) s(n_p) \\ &\quad + \sum_{p=1}^{m} \left[\sum_{q \neq p}^{m} e'(n_p) G(n_p, n_q) s(n_q) \right] \\ &= I_W + I_B \end{aligned} \quad (14)$$

where I_w represents the "within-class" contribution to the Gini Index whereas I_B corresponds to the "between-classes" contribution.

One may observe that the h^{th} element $e'(n_h)G(n_h, n_h)s(n_h)$ of the "within-class" contribution I_w, is equal to

$$\begin{aligned} &e'(n_h) G(n_h, n_h) s(n_h) \\ &= \Big(\underbrace{(1/n) \dots (1/n)}_{n_h \text{ terms}} \Big) \\ &\times \begin{pmatrix} 0 & -1 & \dots & -1 \\ 1 & 0 & \dots & -1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \dots & 0 \end{pmatrix} \begin{pmatrix} s(1, h) \\ s(2, h) \\ \vdots \\ s(n_h, h) \end{pmatrix} \end{aligned} \quad (15)$$

where $s(i, h)$ is the share of individual i, belonging to class h, in society's total income. But expression (15) may also be written, defining $s_{.h}$ as

[3] This is exactly what is implied by the theorem A.5 of Pyatt, Chen and Fei (1980) and by the theorem 8.5 of Kakwani (1980). The share $s_{.i}$ is what Pyatt, Chen and Fei called ϕ_i.

[4] Naturally such a decomposition of inequality remains subject to Shorrocks' (1982) criticism concerning the arbitrariness of decomposition rules.

$\sum_{i=1}^{n_h} s(i, h)$, as

$$e'(n_h)G(n_h, n_h)s(n_h) = (n_h/n)((1/n_h), \dots (1/n_h)) \times \begin{pmatrix} 0 & -1 & \dots & -1 \\ 1 & 0 & \dots & -1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \dots & 0 \end{pmatrix} \times \begin{pmatrix} s(1, h)/s_{.h} \\ s(2, h)/s_{.h} \\ \vdots \\ s(n_h, h)/s_{.h} \end{pmatrix} (s_{.h})$$

$$\leftrightarrow e'(n_h)G(n_h, n_h)s(n_h) = (n_h/n)(s_{.h})I_{Gh} \quad (16)$$

where I_{Gh} is the Gini Index of Inequality within the h^{th} income class. More generally,

$$I_w = \sum_{h=1}^{m} (n_h/n)(s_{.h})I_{Gh}. \quad (17)$$

In other words the "within income classes" contribution I_w to the total Income Inequality I_G, is equal to a weighted sum of the within-classes inequality indices I_{Gh}, the weights being equal to the product of the shares of class h in the total population (n_h/n) and in the total income $(s_{.h})$.

On the other hand the "pq" element of the "between-classes" contribution I_B of expression (14) may be written, assuming for example that $q > p$, as

$$e'(n_p)G(n_p, n_q)s(n_q) = \Big(\underbrace{(1/n), \dots (1/n)}_{n_p \text{ terms}}\Big) \times \begin{pmatrix} -1 & \dots & -1 \\ -1 & \dots & -1 \\ \vdots & & \vdots \\ -1 & \dots & -1 \end{pmatrix} \times \left.\begin{pmatrix} s(1, q) \\ s(2, q) \\ \vdots \\ s(n_q, q) \end{pmatrix}\right\} n_q \text{ terms}$$

$$\leftrightarrow e'(n_p)G(n_p, n_q)s(n_q) = -(n_p/n)(n_q\bar{s}_q) \quad (18)$$

where $\bar{s}_q$ is the mean income share of class q. Similarly, the "qp" element of I_B, assuming $q < p$, is equal to

$$e'(n_q)G(n_q, n_p)s(n_p) = \Big(\underbrace{(1/n), \dots (1/n)}_{n_q \text{ terms}}\Big) \times \begin{pmatrix} 1 & 1 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 1 \end{pmatrix} \left.\begin{pmatrix} s(1, p) \\ s(2, p) \\ \vdots \\ s(n_p, p) \end{pmatrix}\right\} n_p \text{ terms}$$

$$\leftrightarrow e'(n_q)G(n_q, n_p)s(n_p) = (n_q/n)(n_p\bar{s}_p) \quad (19)$$

where $\bar{s}_p$ is the mean income share of class p.

Combining (18) and (19) one derives that

$$I_{B,pq} = e'(n_p)G(n_p, n_q)s(n_q) + e'(n_q)G(n_q, n_p)s(n_p)$$
$$\leftrightarrow I_{B,pq} = -((n_p/n)(n_q\bar{s}_q)) + ((n_q/n)(n_p\bar{s}_p))$$
$$\leftrightarrow I_{B,qp} = ((n_p + n_q)/n) \times (n_p\bar{s}_p + n_q\bar{s}_q)I_{pq} \quad (20)$$

where

$$I_{pq} = ((n_p/(n_p + n_q)), (n_q/(n_p + n_q))) \times \begin{pmatrix} 0 & -1 \\ 1 & -0 \end{pmatrix} \begin{pmatrix} (n_p\bar{s}_p/(n_p\bar{s}_p + n_q\bar{s}_q) \\ (n_q\bar{s}_q/(n_p\bar{s}_p + n_q\bar{s}_q) \end{pmatrix}. \quad (21)$$

Combining (14) and (21) one finally obtains

$$I_B = \sum_{p=1}^{m} \sum_{q>p}^{m} ((n_p + n_q)/n)(n_p\bar{s}_p + n_q\bar{s}_q)I_{pq}. \quad (22)$$

Expression (22) shows that the "between-classes" inequality index I_B is equal to a weighted average of the "between groups p and q" inequality indices I_{pq} defined earlier, the weights being equal to the product of the shares in the total population and income of a group which would include only the income classes p and q.

It can be shown that I_{pq} may also be written as

$$I_{pq} = \Big(\underbrace{\big(1/(n_p + n_q)\big) \dots}_{(n_p + n_q)\text{ identical terms}} \Big) \times \begin{pmatrix} 0 & -1 & \dots & -1 \\ 1 & 0 & \dots & -1 \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \dots & 0 \end{pmatrix} \times \begin{pmatrix} \left.\begin{matrix} (\bar{s}_p/(n_p\bar{s}_p + n_q\bar{s}_q)) \\ \vdots \\ (\bar{s}_p/(n_p s_p + n_q\bar{s}_q)) \end{matrix}\right\} n_p \text{ terms} \\ \left.\begin{matrix} (\bar{s}_q/(n_p\bar{s}_p + n_q\bar{s}_q)) \\ \vdots \\ (\bar{s}_q/(n_p\bar{s}_p + n_q\bar{s}_q)) \end{matrix}\right\} n_q \text{ terms.} \end{pmatrix} \tag{23}$$

Expression (23) indicates that I_{pq}, the "between-classes p and q Gini Index of Inequality," is in fact the Gini Index of a group including only the individuals belonging to the two classes p and q, each individual receiving the average income of the income class to which he belongs.

V. Population Subgroups and the Gini Index: The Case of Overlapping Partitions of the Income Distribution

In the present section an attempt will be made to generalize the results of the previous section. Whereas in the latter it was shown how the Gini Index of Inequality could be broken down into within- and between-income classes contributions, it will now be seen that a decomposition of the Gini Index is possible even in the case of overlapping partitions of the income distributions.

In such a case a third contribution appears, an interaction term, called sometimes a crossover effect (cf. Fei, Ranis and Kuo, (1979)). It will be shown here, firstly, that this interaction term may be very easily computed, secondly that it can be given a very clear and intuitive interpretation.

The problem of inequality decomposition by population subgroups has been analyzed recently in several articles (e.g., Blackorby et al. (1981); Cowell (1980); Cowell and Kuga (1981); Das and Parikh (1982); Shorrocks (1980); Shorrocks (1984)), on the assumption that overall inequality indices may be calculated from the size, mean and inequality values of each population subgroup. This is evidently not possible when one tries to decompose the Gini Index since the ranking of the individuals, in the total population as well as in the subgroups, plays a central role in the computation of this index. Several studies (cf. Bhattacharya and Mahalanobis (1967); Fei, Ranis and Kuo (1979); Mangahas (1975); Mehran (1975); Piesch (1975); Pyatt (1976); Rao (1969)) summarized in Nygård and Sandström (1981) have however succeeded in decomposing the Gini Index into three contributions, a within-groups inequality element, a between-groups component and an interaction term. It will now be shown that the use of the G-matrix should simplify the computations.[5]

Let us assume that k population subgroups are distinguished, that s as before is the ordered column vector of the n individual income shares s_i (with $s_1 \geq s_2 \geq \cdots \geq s_i \geq \cdots \geq s_n$) whereas v is a column vector where the shares are ordered firstly by the size of the population subgroup's average income, secondly within each subgroup by decreasing individual income shares. In other words if $v(j,k)$ is the share in total income of individual j belonging to subgroup k, the two following relations hold:

$$\left(\sum_{j=1}^{n_1} v(j,1)/n_1\right) \geq \cdots, \left(\sum_{j=1}^{n_h} v(j,h)/n_h\right) \dots \geq \left(\sum_{j=1}^{n_k} v(j,k)/n_k\right) \tag{24}$$

and

$$v(1,h) \geq v(2,h) \geq \cdots \geq v(j,h) \geq \cdots \geq v(n_h,h) \quad \forall h \tag{25}$$

where n_h is the population size of subgroup h.

[5] Professor Graham Pyatt indicated justly to me in a private correspondence that footnote 1 in his 1976 paper referred somehow to the existence of the G-matrix since the difference between his matrices A and A' is in fact equal to G. I have to acknowledge that I overlooked his footnote so that the demonstrations given here for the case of non-overlapping as well as overlapping groups could be considered as having been suggested originally in his paper. Pyatt's paper (1976) should be given the paternity for showing the use of matrix algebra in the computation of the Gini Index.

The product $e'Gv$ may be decomposed into two elements, a within-subgroups component which may be written as $\sum_{h=1}^{j} e'(n_h)G(n_h, n_h)v(n_h)$ and a between-groups contribution which may be written as

$$\sum_{h=1}^{k} \sum_{l \neq h}^{k} e'(n_h)G(n_h, n_l)v(n_1).$$

One should notice that this breakdown of v into 2 components only is possible even though some income share $v(j, i)$ of individual j belonging to subgroup i may be lower than some income share $v(t, r)$ of individual t belonging to group r, a possibility which was excluded in the previous section when the income distribution was partitioned in income classes. This is so, firstly because within each subgroup the individuals are ranked by decreasing income share, secondly because the share v_h (v_h is the average income share of group h) used in the computation of the between-groups inequality (cf. expression (22)) are ranked by non-increasing order, a necessary condition for being able to use the G-matrix in the computation of Gini indices.

Having decomposed $e'Gv$ into a between- and a within-groups component, we now compute the value of the product $e'Gd$ where d is a column vector equal to the difference between the vectors s and v. One may therefore write, recalling the linear properties of the G matrix, that

$$e'Gd = e'G(s - v) = e'Gs - e'Gv. \tag{26}$$

The expression $e'Gs$ in (26) is, as was seen earlier, the Gini Index of Inequality. It is well known that the latter is equal to twice the area between the Lorenz curve (cf. Kakwani (1980) for a detailed presentation of the Lorenz curve and of its properties) and the diagonal of a one by one square. The coordinates of the Lorenz curve are the cumulative income shares on one axis and the cumulative population shares on the horizontal axis. It has been shown, however (cf. Atkinson (1980) and Plotnick (1981)), that another curve may be built whose coordinates on the income shares axis would be the cumulative shares corresponding to another ordering of the income shares s_{ji} and that such a curve will always be above the Lorenz curve. In our case this alternative ordering is that corresponding to the shares v_{ji} of vector v. In the context of income redistribution the area between these two curves has been proposed (cf. Atkinson (1980); Plotnick (1981); Plotnick (1982) and Plotnick (1985)) as an index of the horizontal inequity resulting from the reordering of the shares s_{ji} of vector s into the shares v_{ji} of vector v. Since the purpose of this section is to break down the inequality by population subgroup, one may say that the expression $e'Gd = e'G(s - v)$ is a measure of the intensity of the permutations which occur when instead of ranking all the individual shares by decreasing income shares, one ranks them, firstly by decreasing value of the average income of the population subgroup to which they belong, and secondly, within each subgroup, by decreasing individual income share.[6]

We have therefore been able to decompose in a relatively simple way the Gini Inequality Index $e'Gs$ into three components, the within population subgroups inequality, the between population subgroups inequality and an interaction term to which a clear and intuitive interpretation has been given.

VI. Illustrations

Factor Components and the Gini Index

The technique presented in section II has been applied to income data from Taiwan, given in Fei, Ranis and Kuo (1979) and showing firstly the breakdown of Family Income before Tax into income after tax, direct and indirect tax, and second the breakdown of income after tax into housing, educational and other expenditures, and savings. The results obtained using the method described here are almost identical to those given in their tables 6.2 and 6.3 and may be obtained upon request from the author.

The technique presented in section III has been also applied to data from the Retirement History Study, a ten year longitudinal survey of the retirement process by the Social Security Administration. The data which form the basis of our computations are taken from a study of changes in well-being across life by Burkhauser, Butler and Wilkinson (1985). The measure of well being which is used there is what these authors have called comprehensive income. It includes wage and salary earnings as well as the annuitized value of all

[6] It can be shown that $e'G(s - v)$ would be equal, in the context of income redistribution, to twice the value of the area of horizontal inequity defined in Atkinson (1980) and Plotnick (1981).

TABLE 1.—CONTRIBUTION OF VARIOUS INCOME SOURCES TO TOTAL INEQUALITY OF 1969 COMPREHENSIVE INCOMES (RETIREMENT HISTORY STUDY)

Income Source	Contribution to Total Inequality	Within Income Source Pseudo Gini	Total Income Share	Within Income Source Gini Index
Social Security	.016	.094	.170	.095
Private Pension	.037	.414	.089	.414
Wages	.156	.288	.541	.288
Other Wources	.093	.465	.200	.465
Total Income	.303			

TABLE 2.—BREAKDOWN OF THE INCOME INEQUALITY (U.S. CENSUS, 1980)

1) Contribution of the between ethnic group inequality:		0.043
2) Contribution of the within[a] ethnic groups inequality:		0.258
—contribution of whites:	0.254	
—contribution of blacks:	0.003	
—contribution of Asians and Pacific Islanders:	0.000	
—contribution of those of Hispanic origin:	0.001	
3) Contribution of the "Permutation" component:		0.060
Total Inequality (as measured by Gini Index I_G):		0.361

[a] The Gini Index I_G for the various ethnic groups was as follows:

Whites:	$I_G = 0.348$
Blacks:	$I_G = 0.420$
Asians and Pacific Islanders:	$I_G = 0.364$
Hispanic origin:	$I_G = 0.386$

wealth currently held. Using (7), an estimate of the contribution of the various components (Social Security, private pension, wages and other components) to the Gini Index of the Inequality of Comprehensive Incomes was obtained. Estimates of the "Pseudo-Gini Indices" of the various components as well as of the Gini Index of Inequality for each of the four components were also computed. All these estimates are given in table 1, which indicates that wages were in 1969 the main component of the inequality of comprehensive incomes, mainly because their share in total comprehensive income was the highest of all shares (0.542). The second highest contribution came from "other sources" (other than social security, private pensions and wages) not only because this second component had the second highest share in total income (0.200), but also because it was the component with the highest inequality (0.465). It should be stressed that the use of the aggregate data given by Burkhauser, Butler and Wilkinson (1985) implies that one ignores the within-deciles income inequality.

Income Inequality and Population Subgroups

Here again the computation technique was first applied to data given by Fei, Ranis and Kuo (1979). The results obtained were identical to those given in their tables 12.5 and 12.6 and are presented in appendix 1.

The approach described in section V was then applied to data taken from the 1980 U.S. Census and given in Bureau of the Census (1982). There a breakdown of household incomes is given by income and ethnic groups. Nine income groups and four ethnic origins (Whites, Blacks, Asian and Pacific Islanders, Spanish) have been distinguished. Table 2 shows the breakdown of the income inequality into its three components: the between-groups inequality, the within-groups inequality and the interaction term (the permutation component). It should be stressed that here also the within-income group inequality is ignored, given that aggregate data have been used. Table 2 shows that the greatest part of the total inequality results from the within ethnic groups inequality: a contribution of 0.258 out of a total of 0.361. One should notice, however, that the "permutation component" is not negligible since its contribution to total inequality (0.060) is greater than that of the between ethnic groups inequality (0.043). This observation seems to strengthen the case for using the technique based on the G-matrix presented in this paper, since this approach allows one to give an intuitive interpretation to an interaction term which was shown to be empirically significant, and even more important than the between-groups inequality.

APPENDIX I

Income Inequality and Population Subgroups: A Numerical Example from Fei, Ranis and Kuo (1979)

Assume 7 individuals belonging to three different groups including respectively 2, 3 and 2 individuals. The earnings in groups 1, 2 and 3 are respectively (1 and 3), (1, 4 and 7), (6 and 10). Using the computation procedure and the notations presented in section V, one derives:

Gini Inequality Index for the Whole Population

$$e' = ((1/7)(1/7)(1/7)(1/7)(1/7)(1/7)(1/7))$$
$$s' = ((10/32)(7/32)(6/32)(4/32)(3/32)(1/32)(1/32))$$

and

$$I_G = e'G(7,7)s = 84/224.$$

Between Groups Inequality (Intergroup Variation)

$$e' = ((2/7)(3/7)(2/7)),$$
$$s' = ((16/32)(12/32)(4/32))$$

and

$$I_B = e'G(3,3)s = 60/224.$$

Within Groups Inequality (Intragroup Variation)

$$\begin{aligned} I_W = {} & ((1/7)(1/7))G(2,2)((3/32)(1/32)) \\ & + ((1/7)(1/7)(1/7))G((3,3)(7/32)(4/32)(1/32)) \\ & + ((1/7)(1/7)G(2,2)((10/32)(6/32)) = 18/224. \end{aligned}$$

Permutation Intensity (Interaction Term)

The ordering of the average incomes is as follows: group 3 $((10 + 6)/2 = 8)$, group 2 $((1 + 4 + 7)/3 = 4)$ and group 1 $((1 + 3)/2 = 2)$. So the Permutation Intensity I_p which corresponds to twice the crossover effect in Fei, Kuo and Ranis (1979, p. 401) may be computed as $I_P = e'G(7,7)(s - v)$ with $v' = ((10/32)(6/32)(7/32)(4/32)(1/32)(3/32)(1/32))$ so that $e'G(7,7) = 78/224$ and $I_p = ((84 - 78)/224) = 6/224$.

It can be seen that all the result are those obtained by Fei, Ranis and Kuo (1979, page 401), with the exception of two printing mistakes (table 12.6).

REFERENCES

Atkinson, Anthony, "Horizontal Equity and the Distribution of the Tax Burden," in Henry Aaron and Michael Boskin (eds.), *The Economics of Taxation* (Washington, D.C.: The Brookings Institution, 1980).

Berrebi, Z. Moshe, and Jacques Silber, "Income Inequality Indices and Deprivation: A Generalization," *Quarterly Journal of Economics* 99 (Aug. 1985), 807–810.

______, "Regional Differences and the Components of Growth and Inequality Change," *Economics Letters* 25 (1987), 295–298.

Bhattacharya, N., and B. Mahalanobis, "Regional Disparities in Household Consumption in India," *Journal of the American Statistical Association* 62 (Mar. 1967), 143-161.

Blackorby, Charles, David Donaldson, and Maria Auersperg, "A New Procedure for the Measurement of Inequality within and among Population Subgroups," *Canadian Journal of Economics* 14 (Nov. 1981), 665–686.

Bureau of the Census, U.S. 1980 Census of Population and Housing: Provisional Estimates of Social, Economic and Housing Characteristics (Washington, D.C.: U.S. Department of Commerce, Mar. 1982).

Burkhauser, Richard U., J. S. Butler, and James T. Wilkinson, "Estimating Changes in Well-Being Across Life: A Realized vs. Comprehensive Income Approach," in David Martin, and Timothy Smeeding (eds.), *Horizontal Equity, Uncertainty and Economic Well-Being*, Studies in Income and Wealth (Chicago: National Bureau of Economic Research and the University of Chicago Press, 1985).

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

Cowell, Frank, and Kiyoshi Kuga, "Additivity and the Entropy Concept: An Axiomatic Approach to Inequality Measurement," *Journal of Economic Theory* 25 (Aug. 1981), 131–143.

Das, T., and A. Parikh, "Decomposition of Inequality Measures and a Comparative Analysis," *Empirical Economics* 7 (1982), 23–48.

Donaldson, David, and John A. Weymark, "A Single Parameter Generalization of the Gini Indices of Inequality," *Journal of Economic Theory* 22 (Feb. 1980), 67–87.

Fei, John C. H., Gustav Ranis, and Shirley W. Y. Kuo, "Growth and the Family Distribution of Income by Factor Components," *Quarterly Journal of Economics* 92 (Feb. 1978), 17–53.

______, *Growth with Equity, The Taiwan Case* (London: Oxford University Press, 1979).

Kakwani, Nanak C., *Income Inequality and Poverty: Methods of Estimation and Policy Applications* (London: Oxford University Press, 1980).

Lerman, Robert I., and Shlomo Yitzhaki, "A Note on the Calculation and Interpretation of the Gini Index," *Economics Letters* 15 (1984), 363–368.

______, "Income Inequality Effects by Income Source: A New Approach and Applications to the United States,' this REVIEW 67 (Feb. 1985), 151–156.

Mangahas, Mahar, "Income Inequality in the Philippines: A Decomposition Analysis," Population and Employment Working Paper No. 12, International Labor Organization, Geneva, Feb. 1975.

Mehran, Farhad, "A Statistical Analysis of Income Inequality Based on a Decomposition of the Gini Index," *Proceedings of the 40th Session of I.S.I.* (Warsaw, 1975).

Nygård, Fredrik, and Arne Sandström, *Measuring Income Inequality*, Acta Universitatis Stockholmiensis, Stockholm. Studies in Statistics 1 (Stockholm: Almqvist and Wiksell International, 1981).

Piesch, W., *Statistische Konzentrationmasse* (Tübingen: J. C. B. Mohr-Paul Siebeck, 1975).

Plotnick, Robert, "A Measure of Horizontal Inequity," this REVIEW 63 (May 1981), 283–288.

______, "The Concept and Measurement of Horizontal Inequity," *Journal of Public Economics* 17 (Apr. 1982), 373–391.

______, "A Comparison of Measures of Horizontal Inequity," in David Martin and Timothy Smeeding (eds.), *Horizontal Equity, Uncertainty and Economic Well-Being*, Studies in Income and Wealth, 50 (Chicago: National Bureau of Economic Research and the University of Chicago Press, 1985).

Pyatt, Graham, "On the Interpretation and Disaggregation of Gini Coefficients," *Economic Journal* 86 (Mar. 1976), 243–255.

Pyatt, Graham, Chau-Nan Chen, and John C. H. Fei, "The Distribution of Income by Factor Components," *Quarterly Journal of Economics* 95 (Nov. 1980), 451–473.

Rao, V. N., "Two Decompositions of Concentration Ratios," *Journal of the Royal Statistical Society*, Series A 132 (1969), 418–425.

Sen, Amartya, *On Economic Inequality* (London: Oxford University Press, 1973).

Shalit, Haim, "Calculating the Gini Index for Individual Data," *Oxford Bulletin of Economics and Statistics* 47 (May 1985), 185–189.

Shorrocks, Anthony F., "Inequality Decomposition by Factor Components," *Econometrica* 50 (1982), 193–211.

______, "The Impact of Income Components on the Distribution of Family Incomes," *Quarterly Journal of Economics* 97 (1983), 311–326.

______, "Inequality Decomposition by Population Subgroups," *Econometrica* 52 (Nov. 1984), 1369–1385.

Soltow, Lee, "The Distribution of Income Related to Changes in the Distribution of Education, Age and Occupation," this REVIEW 42 (Nov. 1960), 450–453.