Statistica & Applicazioni - Vol. XVI, n. 1, 2018, pp. 31-79

JOINT DECOMPOSITION BY SUBPOPULATIONS AND SOURCES OF THE POINT AND SYNTHETIC GINI INDEXES

Michele Zenga*
Igor Valli*

SUMMARY

The total income Y is the sum of c sources X_j. The Nunits of the population are partitioned into k different subpopulations.Let $\{y_1 < ... < y_h < ... < y_r\}$ the set of the distinct values assumed by Y and $\{n_{1.}, ..., n_{h.}, ..., n_{r.}\}$ are the corresponding frequencies. $G_h(Y)$ and $G(Y) = \sum_{h=1}^{r} G_h(Y) \cdot \frac{n_{h.}}{N}$ are the point and the synthetic Gini (1914) inequality indexes, respectively; $G_h(Y)$ is ''related'' to the area of the h^{th} trapezium of the ''concentration area''. In this paper we have obtained: in the first step the decompositions by sources, by subpopulations and the joint decompositions by subpopulations and sources of $G_h(Y)$; in the second step we have extended the previous decompositions to the synthetic Gini index $G(Y)$. In particular in the first step we have obtained the decomposition $G_h(Y) = \sum_{l=1}^{k} C_{hl.}(Y)$ where $C_{hl.}(Y)$ is the contribution of the subpopulation l to $G_h(Y)$. In addition, $C_{hl.}(Y)$ is decomposed in a within and a between part: $C_{hl.}(Y) = C_{hlW}(Y) + C_{hlB}(Y)$. Moreover for $C_{hl.}(Y)$ we have obtained the following decomposition by sources: $C_{hl.}(Y) = \sum_{j=1}^{c} C_{hl.}(X_j)$, where $C_{hl.}(X_j)$ is the contribution of the source X_j to $C_{hl.}(Y)$. We have also obtained the within $C_{hlW}(X_j)$ and the between part $C_{hlB}(X_j)$ of $C_{hl.}(X_j)$. We have to remark that $G_h(Y)$ is also decomposed in a within and a between part, and in the sum of the k contributions of each sources. The decompositions proposed in this paper are applied to the net disposable income of the 8156 italian households supplied by Bank of Italy (2016) where the households are partitioned in four subpopulations according to the number of family members and the total income is the sum of four sources.

Keywords: Gini Index, ''Two-step'' Approach, Decomposition by Subpopulations, Decomposition by Sources, Joint Decomposition by Subpopulations and Sources.
DOI: 10.26350/999999_000011

1. INTRODUCTION

The total income Y is the sum of c sources X_j: $Y = \sum_{j=1}^{c} X_j$. The N units of a finite population are partitioned into k different subpopulations.

In the case of N values of Y arranged in non decreasing order

$$\left\{0 \leq y_{(1)} \leq \ldots \leq y_{(i)} \leq \ldots \leq y_{(N)} > 0\right\},$$

Zenga (2013) obtained, using the two-step approach, the decompositions by sources of the point and synthetic Gini (1914), Bonferroni (1930) and Zenga M.M. (2007) inequality indexes.

* Dipartimento di Statistica e Metodi Quantitativi - Università Milano-Bicocca - Piazza dell'Ateneo Nuovo, 1 - 20126 MILANO (michele.zenga@unimib.it; ✉ igor.valli@gmail.com).

Let: $Q_{(i)}(Y) = \sum_{t=1}^{i} y_{(t)}, (i = 1, ..., N)$; $T = Q_{(N)}(Y) = \sum_{i=1}^{N} y_{(i)}$;
$M = M(Y) = T/N$; $\bar{M}_{(i)}(Y) = \frac{Q_{(i)}(Y)}{i}$, $(i = 1, ..., N)$.

According to Gini (1914) the relative inequality (concentration), corresponding to the relative frequency $p_{(i)} = \frac{i}{N}$, is given by : $\rho\left(\frac{i}{N}\right) = \frac{p_{(i)} - q_{(i)}}{p_{(i)}}$, $(i = 1, ..., N)$, where $q_{(i)} = \frac{Q_{(i)}(Y)}{T}$ is the ordinate of the Lorenz (1905) curve. The synthetic inequality index $\tilde{G}(Y)$ proposed by Gini (1914) is the weighted mean of $\frac{p_{(i)} - q_{(i)}}{p_{(i)}}$ with weights $p_{(i)}$:

$$\tilde{G}(Y) = \frac{1}{\sum_{i=1}^{N-1} p_{(i)}} \cdot \sum_{i=1}^{N-1} \frac{p_{(i)} - q_{(i)}}{p_{(i)}} \cdot p_{(i)} = \frac{1}{N-1} \cdot \sum_{i=1}^{N-1} 2\left(p_{(i)} - q_{(i)}\right).$$

Another popular expression of $\tilde{G}(Y)$ is given by the ratio: $\tilde{G}(Y) = \frac{\Delta(Y)}{2 \cdot M(Y)}$, where $\Delta(Y) = \frac{1}{N \cdot (N-1)} \cdot \sum_{i=1}^{N} \sum_{j=1}^{N} \left|y_{(i)} - y_{(j)}\right|$ is the Gini mean difference (without replacement) of Y. Recently, Zenga (2016a) obtained the decomposition by subpopulations of the point and synthetic Gini inequality indexes, too. Multiplying both sides of the relation $\tilde{G}(Y) = \frac{1}{N-1} \cdot \sum_{i=1}^{N-1} 2\left(p_{(i)} - q_{(i)}\right)$ by $\frac{N-1}{N}$, we obtain

$$G(Y) = \frac{N-1}{N} \cdot \tilde{G}(Y) = \frac{1}{N} \cdot \sum_{i=1}^{N-1} 2\left(p_{(i)} - q_{(i)}\right) = \frac{1}{N} \cdot \sum_{i=1}^{N} 2\left(p_{(i)} - q_{(i)}\right).$$

In this way $G(Y)$ is the (simple) arithmetic mean of the N point inequality measures $G_{(i)}(Y) = 2\left(p_{(i)} - q_{(i)}\right)$. Note that $G(Y)$ is also equal to the ratio: $G(Y) = \frac{\Delta_R(Y)}{2 \cdot M(Y)}$, where $\Delta_R(Y) = \frac{1}{N^2} \cdot \sum_{i=1}^{N} \sum_{j=1}^{N} \left|y_{(i)} - y_{(j)}\right|$ is the Gini mean difference, with replacement, of Y. Unfortunately, $G_{(i)}(Y)$ may not be constant for units taking the same value of Y. This behaviour of $G_{(i)}(Y)$ is not reasonable in the decomposition by subpopulations because units with the same value of Y may belong to different subpopulations. Zenga (2016a) proposed to overcome this situation by substituting the values of $G_{(i)}(Y)$ corresponding to the $n_{h.}$ units with the same value y_h of Y with their arithmetic mean: $G_h^{\star}(Y) = M\left[G_{(i)}(Y) | Y = y_h\right]$. For this reason, we introduce the set $\{0 \leq y_1 < \ldots < y_h < \ldots < y_r\}$ of the distinct values assumed by the variate Y and the set $\left\{n_{1.}, \ldots, n_{h.}, \ldots, n_{r.}; \sum_{h=1}^{r} n_{h.} = N\right\}$ of the corresponding frequencies. Moreover, we show that $G(Y) = \sum_{h=1}^{r} G_h^{\star}(Y) \cdot \frac{n_{h.}}{N}$ and

$$G_h^{*}(Y) = M\left[G_{(i)}(Y) | Y = y_h\right] = 2(p_{h\cdot} - q_{h\cdot}) - \frac{n_{h\cdot} - 1}{N} \cdot \left(\frac{M(Y) - y_h}{M(Y)}\right)$$
$$= \frac{M(Y) - \bar{M}_{h\cdot}(Y)}{M(Y)} \cdot 2 \cdot p_{h\cdot} - \frac{n_{h\cdot} - 1}{N} \cdot \left(\frac{M(Y) - y_h}{M(Y)}\right),$$

where

$$p_{h.} = \frac{\sum_{t=1}^{h} n_{t.}}{N}; q_{h.} = \frac{\sum_{t=1}^{h} y_t \cdot n_{t.}}{T}; \bar{M}_{h.}(Y) = \frac{\sum_{t=1}^{h} y_t \cdot n_{t.}}{\sum_{t=1}^{h} n_{t.}}.$$

The decomposition by subpopulations of $G_h^\star(Y)$ is essentially based (first step) on the decompositions in $k \times k$ additive terms of the difference $(M(Y) - \bar{M}_{h.}(Y))$ and of the difference $(M(Y) - y_h)$. Then, the decomposition by subpopulations of $G(Y)$ is obtained (second step) putting the decomposition of $G_h^\star(Y)$ in the relation $G(Y) = \sum_{h=1}^{r} G_h^\star(Y) \cdot \frac{n_{h.}}{N}$.

The two-step approach is used in the present paper to obtain the joint decomposition by sources and by subpopulations of the Gini inequality index, too. We remark that, in the present paper, the value of $G(Y)$ is obtained by the following expression related to the Lorenz curve: $G(Y) = \frac{CA}{1/2} = \sum_{h=1}^{r} G_h(Y) \cdot \frac{n_{h.}}{N}$, where: CA is the so called ''concentrantion'' area whose value is evaluated by

$$CA = \sum_{h=1}^{r} K_h = \sum_{h=1}^{r} \frac{1}{2} [(p_{h.} - q_{h.}) + (p_{h-1.} - q_{h-1.})]. \frac{n_{h.}}{N},$$

and

$$G_h(Y) = \frac{2 \cdot K_h}{(n_{h.})/N} = (p_{h.} - q_{h.}) + (p_{h-1.} - q_{h-1.}) = 2(p_{h.} - q_{h.}) - \frac{n_{h.}}{N} \cdot \left(\frac{M(Y) - y_h}{M(Y)} \right).$$

Finally,

$$G(Y) = \sum_{h=1}^{r} G_h^\star(Y) \cdot \frac{n_{h.}}{N} = \sum_{h=1}^{r} G_h(Y) \cdot \frac{n_{h.}}{N} + \frac{1}{N \cdot M} \cdot \sum_{h=1}^{r} (M(Y) - y_h) \cdot \frac{n_{h.}}{N}.$$

From a property of the arithmetic mean: $\sum_{h=1}^{r} (M(Y) - y_h) \cdot \frac{n_{h.}}{N} = 0$. Consequently, $G(Y) = \sum_{h=1}^{r} G_h^\star(Y) \cdot \frac{n_{h.}}{N} = \sum_{h=1}^{r} G_h(Y) \cdot \frac{n_{h.}}{N}$.

The remainder of this paper is organaized as follows. Section 2 introduces the definitions and notation necessary, in the case of the frequency distribution framework of the total income Y, to describe the point and the synthetic Gini inequality measures. It is worth to remember that, in the present paper, the values of $G_h(Y)$ are related to the so called ''concentration area'' of the Lorenz curve. Section 3 introduces the means $M(X_j)$ and the lower means $\bar{M}_{h.}(X_j)$ of the components of the total income $Y = \sum_{j=1}^{c} X_j$. Then, by the use of the relation $M(Y) = \sum_{j=1}^{c} M(X_j)$, the following decomposition by sources of $G_h(Y)$ has been obtained: $G_h(Y) = \sum_{j=1}^{c} C_h(X_j)$, where $C_h(X_j)$ is the additive contribution of X_j to $G_h(Y)$. Section 3.1 provides a numerical illustration of the decompositon by sources of $G_h(Y)$ and of $G(Y)$. The bivariate distribution of the N units according to k different subpopulations and the r distinct values of the total income Y plays a significant role in the decomposition by subpopulations. This distribution is illustrated in Section 4. Moreover, this section provides for each subpopulation g: the mean $M_g(Y)$; the lower mean $\bar{M}_{hg}(Y)$; the relative frequency $\frac{n_{.g}}{N}$, where $n_{.g}$ is the size of subpopulation g; the relative frequency $p(g|h)$ of subpopulation g in the lower group; the relative frequency $f(g|h)$ of subpopulation g in the group of $n_{h.}$ units with $Y = y_h$. Then, by the use of the relations $M(Y) = \sum_{g=1}^{k} M_g(Y) \cdot \frac{n_{.g}}{N}$, $\bar{M}_{h.}(Y) = \sum_{l=1}^{k} \bar{M}_{hl}(Y) \cdot p(l|h)$ and $y_h = \sum_{l=1}^{k} y_h \cdot f(l|h)$, the following $k \times k$ additive decomposition by subpopulations of $G_h(Y)$ is obtained: $G_h(Y) = \sum_{l=1}^{k} \sum_{g=1}^{k} C_{hlg}(Y)$, where $C_{hlg}(Y)$ is the contribution to $G_h(Y)$ that derives from the comparisons of $\bar{M}_{hl}(Y)$ and y_h w.r.t. $M_g(Y)$. Section 4.1 gives a numerical illustration of the decomposition by subpopulations obtained in Section 4. Section 5 gives for each subpopulation g: the cell means $M_{hg}(X_j)$ and $M_{hg}(Y)$; the lower means $\bar{M}_{hg}(X_j)$ and $\bar{M}_{hg}(Y)$; and the subpopulation means $M_g(X_j)$ and $M_g(Y)$. Then, by the use of the relations $M_g(Y) = \sum_{j=1}^{c} M_g(X_j)$, $\bar{M}_{hl}(Y) = \sum_{j=1}^{c} \bar{M}_{hl}(X_j)$ and $y_h = M_{hl}(Y) = \sum_{j=1}^{c} M_{hl}(X_j)$ in the decomposition by subpopulations of $G_h(Y)$ reported in Section 4, the basic $k \times k \times c$ joint decomposition by subpopulations and sources of $G_h(Y)$ is obtained: $G_h(Y) = \sum_{l=1}^{k} \sum_{g=1}^{k} \sum_{j=1}^{c} C_{hlg}(X_j)$. For the example at hand, Section 5.2 reports the calculation for $C_{1lg}(X_j)$, $j = 1, 2, 3$. Section 6 gives the decompositions by sources of the point and the synthetic indexes of each subpopulation l. Section 7 provides an application to the net disposable income of the Italian households supplied by the 2014 Central Bank of Italy sample survey on household income and wealth (Bank of Italy 2016). Section 8 is devoted to the conclusions

2. DEFINITIONS AND NOTATION

Let, $X_1, \ldots, X_j, \ldots, X_c$ be non-negative variates (income sources) observable on each of the N units of the population, and $Y = \sum_{j=1}^{c} X_j$ be the total income. Let: $\{0 \leq y_1 < \ldots < y_h < \ldots < y_r\}$ be the set of the distinct values assumed by the variate Y and $\{n_{1.}, \ldots, n_{h.}, \ldots, n_{r.}; \sum_{h=1}^{r} n_{h.} = N\}$ be the corresponding frequencies. At each y_h we can consider the lower group $\{Y \leq y_h\}$, including the first $P_{h.} = \sum_{t=1}^{h} n_{t.}$ units with cumulative income $Q_{h.}(Y) = \sum_{t=1}^{h} y_t \cdot n_{t.}$. Let,

$$\bar{M}_{h.}(Y) = \frac{Q_{h.}(Y)}{P_{h.}}, \qquad h = 1, \ldots, r, \tag{1}$$

be the arithmetic mean (lower mean) of the lower group, and

$$\bar{M}_{r.}(Y) = M(Y) = \frac{T}{N} \tag{2}$$

be the arithmetic mean of the overall population. Note that, in formula (2), $T = Q_{r.}(Y)$. The ratios $p_{h.} = \frac{P_{h.}}{N}$ and $q_{h.} = \frac{Q_{h.}}{T}$, $h = 1, ..., r$ are respectively, the (cumulative) relative frequency and the income share of the lower group. Moreover, $\{(p_{h.}, q_{h.}) : h = 0, 1, \ldots, r; (p_{0.} = 0,\ q_{0.} = 0)\}$ is the set of the coordinates of the Lorenz (1905) curve.

The following $N = 10$ ordered values $\{y_{(i)} : 2, 2, 8, 24, 29, 37, 37, 37, 62, 62\}$ are utilized in Table 1 for the calculation of $p_{h.}$, $q_{h.}$ and $\bar{M}_{h.}(Y)$. Table 1 shows that: $M(Y) = \bar{M}_{6.} = 30$.

Figure 1 reports the graph of the Lorenz curve for the distribution presented in Table 1. In this graph, the egalitarian line (OD) and the concentration polygonal chain (OBEFSWD), enclose the so called ‘‘concentration area’’ CA. The value of CA is given by:

$$CA = \sum_{h=1}^{r} K_h = \sum_{h=1}^{r} \frac{1}{2} [(p_{h.} - q_{h.}) + (p_{h-1.} - q_{h-1.})]. \frac{n_{h.}}{N}, \quad (3)$$

where K_h is the area of the h^{th} trapezium $(FSKV)$. In Figure 1,

$$K_4 = \frac{1}{2} [(p_{4.} - q_{4.}) + (p_{3.} - q_{3.})]. \frac{n_{4.}}{N} = 0.02816.$$

TABLE 1. - *Calculation of: $P_{h.}$, $Q_{h.}$, $\bar{M}_{h.}, p_{h.}$ and q_h*

h	y_h	$n_{h.}$	$P_{h.}$	$y_h \cdot n_{h.}$	$Q_{h.}$	$\bar{M}_{h.}$	$p_{h.}$	$q_{h.}$
1	2	2	2	4	4	2	0.2	0.013
2	8	1	3	8	12	4	0.3	0.04
3	24	1	4	24	36	9	0.4	0.12
4	29	1	5	29	65	13	0.5	0.216
5	37	3	8	111	176	22	0.8	0.586
6	62	2	10	124	300	30	1	1
		10		300				

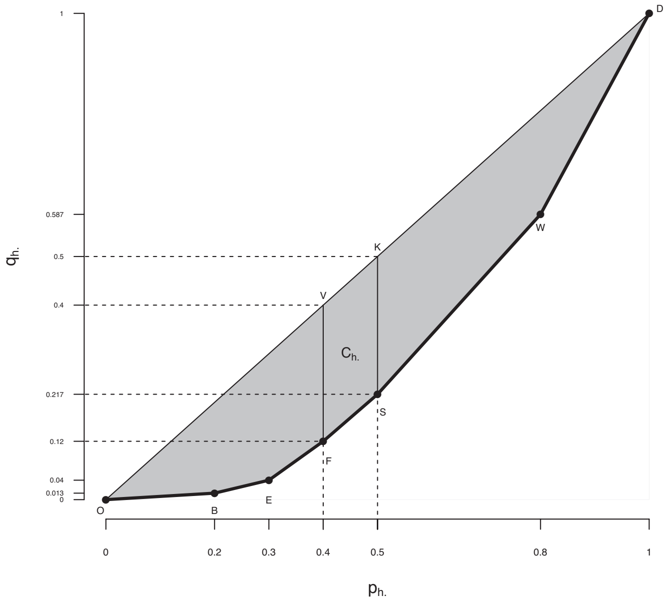

FIGURE 1. - *Graph of the Lorenz curve*

The synthetic Gini inequality index $G(Y)$ can be evaluated by:

$$G(Y) = \frac{CA}{1/2} = \sum_{h=1}^{r} G_h(Y) \cdot \frac{n_{h.}}{N} \quad (4)$$

where $1/2$ is the area of the triangle (ODZ), $G_h(Y) \cdot \frac{n_{h.}}{N} = 2 \cdot K_h$, and

$$G_h(Y) = \frac{2K_h}{\frac{n_{h.}}{N}} = [(p_{h.} - q_{h.}) + (p_{h-1.} - q_{h-1.})] \tag{5}$$

is the sum of the segments $\overline{KS}$ and $\overline{VF}$ of the trapezium $(FSKV)$. Using in (5) the relations $p_{h-1.} = p_{h.} - \frac{n_{h.}}{N}$ and $q_{h-1.} = q_{h.} - \frac{y_h}{M} \cdot \frac{n_{h.}}{N}$, $(h = 1, \ldots, r)$, gives

$$G_h = \left\{ 2(p_{h.} - q_{h.}) - \frac{n_{h.}}{N} \cdot \left(\frac{M(Y) - y_h}{M(Y)} \right) \right\}. \tag{6}$$

Then, $G(Y)$ can be also evaluated by

$$G(Y) = \sum_{h=1}^{r} \left\{ 2(p_{h.} - q_{h.}) - \frac{n_{h.}}{N} \cdot \left(\frac{M(Y) - y_h}{M(Y)} \right) \right\} \cdot \frac{n_{h.}}{N}. \tag{7}$$

From (1) and (2) it derives that:

$$q_{h.} = \frac{Q_{h.}}{T} = \frac{\bar{M}_{h.}(Y)}{M(Y)} \cdot \frac{P_{h.}}{N} = \frac{\bar{M}_{h.}(Y)}{M(Y)} \cdot p_{h.}. \tag{8}$$

Thus,

$$p_{h.} - q_{h.} = \frac{M(Y) - \bar{M}_{h.}(Y)}{M(Y)} \cdot p_{h.} \tag{9}$$

and

$$\frac{p_{h.} - q_{h.}}{p_{h.}} = \frac{M(Y) - \bar{M}_{h.}(Y)}{M(Y)}. \tag{10}$$

Now, using (9) and (10) in (6) gives,

$$G_h = \left\{ \frac{M(Y) - \bar{M}_{h.}(Y)}{M(Y)} \cdot p_{h.} \cdot 2 - \frac{n_{h.}}{N} \cdot \frac{M(Y) - y_h}{M(Y)} \right\}. \tag{11}$$

In conclusion, the Gini inequality index $G(Y)$ can also be evaluated by

$$G(Y) = \sum_{h=1}^{r} \left\{ \frac{M(Y) - \bar{M}_{h.}(Y)}{M(Y)} \cdot p_{h.} \cdot 2 - \frac{n_{h.}}{N} \cdot \left(\frac{M(Y) - y_h}{M(Y)} \right) \right\} \cdot \frac{n_{h.}}{N}. \tag{12}$$

The equivalent expressions (5), (6) and (11) can be assumed as a ‘‘Gini point measure of inequality’’.

In this paper, the decompositions by sources, by subpopulations and the joint decomposition of the Gini inequality index, are obtained using (11) and (12). Table 2 reports the calculation of these latter expressions.

The value of $G(Y) = \sum_{h=1}^{r} G_h(Y) \cdot \frac{n_{h.}}{N}$ can be interpreted as the sum of the areas of r rectangles, with basis $\frac{n_{h.}}{N}$ and height $G_h(Y)$, $(h = 1, \ldots, r)$. To draw the inequality diagram G_h, it is necessary, first of all, to obtain the r points of coordinates

$(p_{h.}, G_h)$. Then, we obtain r rectangles by the following procedure: the first rectangle has abscissas in the interval $[0, p_{1.}]$and ordinates in the interval $[0, G_1]$. The h^{th} rectangle, $h = 2, ..., r$, has abscissas in the interval $[p_{h-1.}, p_{h.}]$ and ordinates in the interval $[0, G_h]$. Figure 2 reports the Graph of G_h.

TABLE 2. - *Calculation of: G_h and $G(Y)$*

h	y_h	$\bar{M}_{h.}(Y)$	a $p_{h.}$	b $\frac{n_{h.}}{N}$	c $\frac{M(Y)-\bar{M}_{h.}(Y)}{M}$	$d = c \cdot 2a$	e $\frac{M(Y)-y_h}{M(Y)}$	$f = b \cdot e$	$G_h = d - f$	$G_h \cdot \frac{n_{h.}}{N}$
1	2	2	0.20	0.2	0.933	0.373	0.933	0.186	0.186	0.0373
2	8	4	0.3	0.1	0.866	0.520	0.733	0.073	0.446	0.0446
3	24	9	0.4	0.1	0.700	0.560	0.200	0.02	0.54	0.0540
4	29	13	0.5	0.1	0.566	0.566	0.033	0.003	0.563	0.0563
5	37	22	0.8	0.3	0.266	0.426	-0.233	-0.07	0.496	0.1490
6	62	30	1	0.2	0.000	0.000	-1.066	-0.213	0.213	0.0426
										$G(Y) =$ 0.3840

3. DECOMPOSITION BY SOURCES

The sums of the values of X_j and of Y, observable on each of the $n_{h.}$ units with $Y = y_h$, are denoted with $S_{h.}(X_j)$ and $S_{h.}(Y)$ respectively; note that $S_{h.}(Y) = y_h \cdot n_{h.}$. The corresponding means are:

$$M_{h.}(X_j) = \frac{S_{h.}(X_j)}{n_{h.}} \tag{13}$$

and

$$M_{h.}(Y) = \frac{S_{h.}(Y)}{n_{h.}} = \frac{y_h \cdot n_{h.}}{n_{h.}} = y_{h.} \,. \tag{14}$$

Now, from the relation $Y = \sum_{j=1}^{c} X_j$ and a popular property of the arithmetic mean we can write:

$$y_h = M_h.(Y) = \sum_{j=1}^{c} M_h.(X_j) \qquad (h = 1, \ldots, r). \tag{15}$$

The sum of the values of X_j, observable on each of the $P_{h.}$ units of the lower group $\{Y \leq y_h\}$, is denoted with

$$Q_{h.}(X_j) = \sum_{t=1}^{h} S_{t.}(X_j) \tag{16}$$

and the corresponding mean is

$$\bar{M}_{h.}(X_j) = \frac{Q_{h.}(X_j)}{P_{h.}}. \tag{17}$$

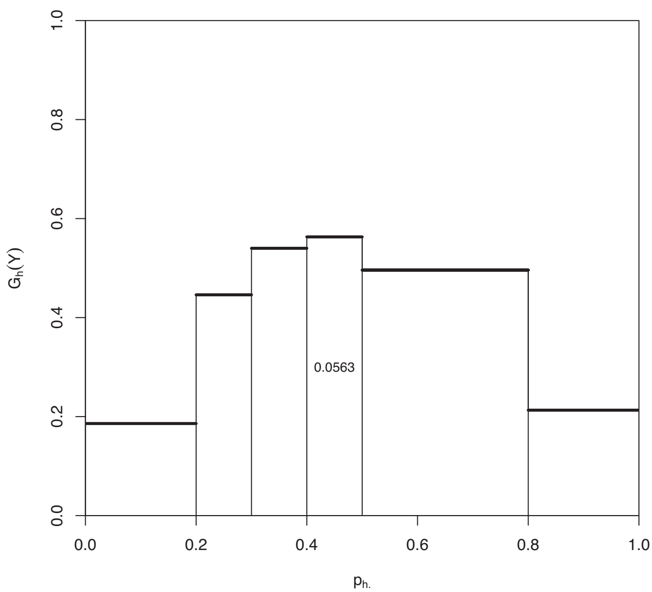

FIGURE 2. - *Gini inequality "curve"* $G_h(Y)$

The sum of the values of Y in the lower group $Q_{h.}(Y) = \sum_{t=1}^{h} y_t \cdot n_{t.}$ and the corresponding lower mean $\bar{M}_{h.}(Y) = \frac{Q_{h.}(Y)}{P_{h.}}$, has been already introduced in Section 2. Obviously,

$$Q_{h.}(Y) = \sum_{j=1}^{c} Q_{h.}(X_j), \tag{18}$$

and

$$\bar{M}_{h.}(Y) = \sum_{j=1}^{c} \bar{M}_{h.}(X_j) \qquad (h = 1, ..., r). \tag{19}$$

For $h = r$, $Q_{r.}(X_j) = T(X_j)$ is the sum of all the N values of X_j, and

$$\bar{M}_{r.}(X_j) = \frac{Q_{r.}(X_j)}{N} = \frac{T(X_j)}{N} = M(X_j). \tag{20}$$

Thus, for $h = r$ the relation (19) gives

$$M(Y) = \sum_{j=1}^{c} M(X_j). \tag{21}$$

In order to obtain the decomposition by sources of $G_h(Y)$ we decompose, first of all, the differences $[M(Y) - \bar{M}_{h.}(Y)]$ and $[M(Y) - y_h]$. Thus, from 21 and (19) we have

$$\begin{aligned} M(Y) - \bar{M}_{h.}(Y) &= \sum_{j=1}^{c} M(X_j) - \sum_{j=1}^{c} \bar{M}_{h.}(X_j) \\ &= \sum_{j=1}^{c} \{M(X_j) - \bar{M}_{h.}(X_j)\}. \end{aligned} \tag{22}$$

Analogously, from (21) and (15) we have

$$M(Y) - y_{h.} = \sum_{j=1}^{c} M(X_j) - \sum_{j=1}^{c} M_{h.}(X_j) = \sum_{j=1}^{c} \{M(X_j) - M_{h.}(X_j)\}. \tag{23}$$

Now, putting (22) and (23) in (11) gives (first step) the following additive decomposition by sources of $G_h(Y)$.

$$G_h(Y) = \sum_{j=1}^{c} C_h(X_j) \tag{24}$$

where

$$C_h(X_j) = \frac{M(X_j) - \bar{M}_{h.}(X_j)}{M(Y)} \cdot p_{h.} \cdot 2 - \frac{M(X_j) - M_{h.}(X_j)}{M(Y)} \cdot \frac{n_{h.}}{N} \tag{25}$$

is the contribution of X_j to $G_h(Y)$. Finally (second step), putting (24) and (25) in (12), the following additive decomposition by sources of $G(Y)$ is obtained:

$$G(Y) = \sum_{h=1}^{r} \left\{ \sum_{j=1}^{c} C_h(X_j) \right\} \cdot \frac{n_{h.}}{N} = \sum_{j=1}^{c} \sum_{h=1}^{r} C_h(X_j) \cdot \frac{n_{h.}}{N} = \sum_{j=1}^{c} C(X_j). \tag{26}$$

In (26)

$$C(X_j) = \sum_{h=1}^{r} C_h(X_j) \cdot \frac{n_{h.}}{N} \tag{27}$$

is the contribution of X_j to the synthetic index $G(Y)$. Note that $C(X_j)$ is the arithmetic mean of the r contributions of X_j to the point indexes $G_h(Y)$.

3.1 *Numerical illustration of the decomposition by sources of $G_h(Y)$ and $G(Y)$*

For the example of Section 2, let us suppose that $c = 3$ and that the values of $S_{h.}(X_j)$ are those reported in Table 3.

TABLE 3. - *Calculation of $M_{h.}(X_j), M_{h.}(Y), M(X_j)$ and $M(Y)$*

		$S_{h.}(\cdot)$				$M_{h.}(\cdot)$			
h	$n_{h.}$	X_1	X_2	X_3	Y	X_1	X_2	X_3	Y
1	2	2	0	2	4	1	0	1	2
2	1	1	3	4	8	1	3	4	8
3	1	5	9	10	24	5	9	10	24
4	1	4	15	10	29	4	15	10	29
5	3	30	39	42	111	10	13	14	37
6	2	28	50	46	124	14	25	23	62
Tot.	10	70	116	114	300	$M(X_1) = 7$	$M(X_2) = 11.6$	$M(X_3) = 11.4$	$M(Y) = 30$

Table 4 reports the cumulative incomes $Q_{h.}(X_j)$ and $Q_{h.}(Y)$, and the lower means $\bar{M}_{h.}(X_j)$ and $\bar{M}_{h.}(Y)$.

TABLE 4. - *Cumulative incomes $Q_{h.}(X_j)$ and $Q_{h.}(Y)$, and lower means $\bar{M}_{h.}(X_j)$ and $\bar{M}_{h.}(Y)$*

	$Q_{h.}(\cdot)$					$\bar{M}_{h.}(\cdot)$			
h	X_1	X_2	X_3	Y	$P_{h.}$	X_1	X_2	X_3	Y
1	2	0	2	4	2	1	0	1	2
2	3	3	6	12	3	1	1	2	4
3	8	12	16	36	4	2	3	4	9
4	12	27	26	65	5	2.4	5.4	5.2	13
5	42	66	68	176	8	5.25	8.25	8.5	22
6	70	116	114	300	10	7	11.6	11.4	30

For the calculation of $C_h(X_j)$ the following relations are useful:

$$C_h(X_j) = V_h(X_j) \cdot 2p_{h.} - A_h(X_j) \cdot \frac{n_{h.}}{N}, \tag{28}$$

where

$$V_h(X_j) = \frac{M(X_j) - \bar{M}_{h.}(X_j)}{M(Y)} \tag{29}$$

and

$$A_h(X_j) = \frac{M(X_j) - M_{h.}(X_j)}{M(Y)}. \tag{30}$$

The calculation of the contributions $C_h(X_j)$ and $C(X_j)$ are reported in the Tables 5 and 6, respectively.

TABLE 5. -*Calculation of* $C_h(X_j) = V_h(X_j) \cdot 2p_{h.} - A_h(X_j) \cdot \frac{n_{h.}}{N}$

	X_1			X_2			X_3		
h	$V_h(\cdot)$	$A_h(\cdot)$	$C_h(\cdot)$	$V_h(\cdot)$	$A_h(\cdot)$	$C_h(\cdot)$	$V_h(\cdot)$	$A_h \frac{n_{h.}}{N} \cdot (\cdot)$	$C_h(\cdot)$
1	0.2000	0.200	0.040	0.3866	0.3866	0.0773	0.346	0.346	0.0693
2	0.2000	0.200	0.100	0.3530	0.2860	0.1833	0.313	0.246	0.1630
3	0.1660	0.066	0.126	0.2860	0.0860	0.2206	0.246	0.046	0.1926
4	0.1530	0.100	0.143	0.2060	-0.1130	0.2180	0.206	0.046	0.2020
5	0.0583	-0.100	0.123	0.1116	-0.0460	0.1926	0.096	-0.086	0.1806
6	0.0000	-0.233	0.046	0.0	-0.4460	0.0893	0.0	-0.386	0.0773

TABLE 6. - *Calculation of* $C(X_j) = \sum_{h=1}^{r} C_h(X_j) \cdot \frac{n_{h.}}{N}$

h	$\frac{n_{h.}}{N}$	$C_h(X_1) \cdot \frac{n_{h.}}{N}$	$C_h(X_2) \cdot \frac{n_{h.}}{N}$	$C_h(X_3) \cdot \frac{n_{h.}}{N}$	Tot.
1	0.2	0.008	0.01546	0.01386	0.0373
2	0.1	0.0100	0.01833	0.01633	0.0446
3	0.1	0.0126	0.02206	0.01926	0.0540
4	0.1	0.0143	0.02180	0.0202	0.0563
5	0.3	0.0370	0.05780	0.0542	0.1490
6	0.2	0.0093	0.01786	0.0154	0.0426
Tot.		0.0913= $C(X_1)$	0.1533= $C(X_2)$	0.1393= $C(X_3)$	0.384= $G(Y)$

4. DECOMPOSITION BY SUBPOPULATIONS OF THE POINT AND SYNTHETIC GINI INDEX

We introduce now the appropiate definitions and notation used in the decomposition by subpopulations proposed in this paper for $G_h(Y)$ and $G(Y)$. In particular we consider the bivariate distribution of the N units according to k different subpopulations and the r distinct values $\{0 \leq y_1 < ... < y_h < ... < y_r\}$ assumed by the variate Y over these subpopulations. It is possible to report the whole $r \times k$ bivariate distribution of the N units as in Table 7, where: n_{hg} denotes the frequency of y_h in the subpopulation g, $n_{h.} = \sum_{g=1}^{k} n_{hg}$ is the frequency of y_h in the whole population, and $n_{.g} = \sum_{h=1}^{r} n_{hg}$ is the size of the subpopulation g.

TABLE 7. - *Bivariate $r \times k$ distribution of the whole population partitioned into k subpopulations*

	Subpopulation					
	1	...	g	...	k	Tot
y_1	n_{11}	...	n_{1g}	...	n_{1k}	$n_{1.}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
y_h	n_{h1}	...	n_{hg}	...	n_{hk}	$n_{h.}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
y_r	n_{r1}	...	n_{rg}	...	n_{rk}	$n_{r.}$
Tot	$n_{.1}$	...	$n_{.g}$	...	$n_{.k}$	N

For the distribution $\{(y_h, n_{hg}) : h = 1, \ldots, r\}$ of the subpopulation g, let

$$P_{hg} = P_{hg}(Y) = \sum_{t=1}^{h} n_{tg} \qquad h = 1, \ldots, r \tag{31}$$

$$Q_{hg}(Y) = \sum_{t=1}^{h} y_t \cdot n_{tg} \qquad h = 1, \ldots, r \tag{32}$$

$$T_g = Q_{rg}(Y) = \sum_{h=1}^{r} y_h \cdot n_{hg}, \tag{33}$$

$$M_g = M_g(Y) = T_g / n_{.g}. \tag{34}$$

In (34), $M_g(Y)$ denotes the mean of subpopulation g. For the same subpopulation g, let

$$o(g) = \min h : \, n_{hg} > 0, \tag{35}$$

and define the lower mean $\bar{M}_{hg}(Y)$ as follows:

$$\bar{M}_{hg}(Y) = \begin{cases} y_{o(g)} & \text{for } h < o(g) \\ Q_{hg}(Y)/P_{hg} & \text{for } \; h \geq o(g) \end{cases} \tag{36}$$

The mean $M(Y)$ is related to the k means $M_g(Y)$ by the relation

$$M(Y) = \sum_{g=1}^{k} M_g(Y) \cdot \frac{n_{\cdot g}}{N}, \tag{37}$$

and $\bar{M}_{h.}(Y)$ is related to the k means $\bar{M}_{hg}(Y)$ by the relations

$$\bar{M}_{h.}(Y) = \sum_{l=1}^{k} \bar{M}_{hl}(Y) \cdot p(l|h), \tag{38}$$

where

$$p(l|h) = \frac{P_{hl}}{P_{h.}} \qquad (h = 1, ..., r; l = 1, ..., k) \tag{39}$$

is the relative frequency of subpopulation l in the lower group $\{Y \leq y_h\}$.

Note that $\sum_{l=1}^{k} p(l|h) = 1$. The value y_h can be written as follows

$$y_h = \sum_{l=1}^{k} y_h \cdot f(l|h) \tag{40}$$

where

$$f(l|h) = \frac{n_{hl}}{n_{h.}}, \qquad (h = 1, ..., r; l = 1, ..., k) \tag{41}$$

is the relative frequency of the subpopulation l in the group of the $n_{h.}$ units with $Y = y_{h.}$. Zenga (2016a), using (37) and (38), obtained for $[M(Y) - \bar{M}_{h.}(Y)]$ the following $k \times k$ additive decomposition:

$$[M(Y) - \bar{M}_{h.}(Y)] = \sum_{l=1}^{k} \sum_{g=1}^{k} \left[M_g(Y) - \bar{M}_{hl}(Y)\right] \cdot \frac{n_{\cdot g}}{N} \cdot p(l|h), \tag{42}$$

and, using (37) and (40), obtained for $[M - y_h]$ the following $k \times k$ additive decomposition:

$$[M(Y) - y_h] = \sum_{l=1}^{k} \sum_{g=1}^{k} \left[M_g(Y) - y_h\right] \cdot \frac{n_{\cdot g}}{N} \cdot f(l|h). \tag{43}$$

Now, putting these two decompositions in (11) gives (first-step) the following $k \times k$ additive decomposition by subpopulations of $G_h(Y)$:

$$G_h(Y) = \sum_{l=1}^{k} \sum_{g=1}^{k} C_{hlg}(Y), \tag{44}$$

where

$$C_{hlg}(Y) = \frac{M_g(Y) - \bar{M}_{hl}(Y)}{M(Y)} \cdot \frac{n_{\cdot g}}{N} \cdot p(l|h) \cdot 2p_{h.} - \frac{M_g(Y) - y_h}{M(Y)} \cdot \frac{n_{\cdot g}}{N} f(l|h) \cdot \frac{n_{h.}}{N} \tag{45}$$

is the contribution to $G_h(Y)$ that derives from the comparisons of $\bar{M}_{hl}(Y)$ and y_h w.r.t. $M_g(Y)$.

Now, putting (40) in (12) gives (second step) the following $k \times k$ additive decomposition by subpopulations of $G(Y)$:

$$G(Y) = \sum_{h=1}^{r} \left\{ \sum_{l=1}^{k} \sum_{g=1}^{k} C_{hlg}(Y) \right\} \cdot \frac{n_{h.}}{N} = \sum_{l=1}^{k} \sum_{g=1}^{k} C_{.lg}(Y), \tag{46}$$

where

$$C_{.lg}(Y) = \sum_{h=1}^{r} C_{hlg}(Y) \cdot \frac{n_{h.}}{N}. \tag{47}$$

4.1 *Numerical illustration of the decomposition by subpopulations of the Gini point inequality measure $G_h(Y)$*

The results of Section 4 are illustrated by the 6×3 bivariate distribution, reported in Table 8, with $N = 10$ units, $k = 3$ subpopulations and $r = 6$ distinct values of Y.

TABLE 8. - *Joint frequencies n_{hg}, total frequencies $n_{.g}$ and $n_{h.}$, cumulative frequencies P_{hg} and $P_{h.}$, and cumulative incomes Q_{hg} and $Q_{h.}$*

		g			tot	g			tot	g			tot
		1	2	3		1	2	2		1	2	3	
h	y_h	n_{h1}	n_{h2}	n_{h3}	$n_{h.}$	P_{h1}	P_{h2}	P_{h3}	$P_{h.}$	Q_{h1}	Q_{h2}	Q_{h3}	$Q_{h.}$
1	2	1	0	1	2	1	0	1	2	2	-	2	4
2	8	0	1	0	1	1	1	1	3	2	8	2	12
3	24	1	0	0	1	2	1	1	4	26	8	2	36
4	29	0	0	1	1	2	1	2	5	26	8	31	65
5	37	2	1	0	3	4	2	2	8	100	45	31	176
6	62	1	0	1	2	5	2	3	10	162	45	93	300
total $n_{.g}$		5	2	3	10								

Table 9 reports all the values necessary for the decomposition by subpopulations of $G_h(Y)$ proposed in the present paper.

$$G_h(Y) = \sum_{l=1}^{k} \sum_{g=1}^{k} C_{hlg}(Y),$$

where

$$C_{hlg}(Y) = \frac{M_g(Y) - \bar{M}_{hl}(Y)}{M(Y)} \cdot \frac{n_{.g}}{N} \cdot p(l|h) \cdot 2p_{h.} - \frac{M_g(Y) - M_{hl}(Y)}{M(Y)} \cdot \frac{n_{.g}}{N} \cdot f(l|h) \cdot \frac{n_{h.}}{N}.$$

TABLE 9. - *Lower means $\bar{M}_{hg}(Y)$ and $\bar{M}_{h.}(Y)$, and relative frequencies $p(l|h) = \frac{P_{hl}}{P_{h.}}$ and $f(l|h) = \frac{n_{hl}}{n_{h.}}$*

		g			$\bar{M}_{h.}$	l			l		
		1	2	3		1	2	3	1	2	3
$o(g)$		1	2	1							
h	y_h	$\bar{M}_{h1}$	$\bar{M}_{h2}$	$\bar{M}_{h3}$		$\frac{P_{h1}}{P_{h.}}$	$\frac{P_{h2}}{P_{h.}}$	$\frac{P_{h3}}{P_{h.}}$	$\frac{n_{h1}}{n_{h.}}$	$\frac{n_{h2}}{n_{h.}}$	$\frac{n_{h3}}{n_{h.}}$
1	2	2	8	2	2	0.5	0	0.5	0.5	0	0.5
2	8	2	8	2	4	0.33	0.33	0.33	0	1	0
3	24	13	8	2	9	0.5	0.25	0.25	1	0	0
4	29	13	8	15.5	13	0.4	0.2	0.4	0	0	1
5	37	25	22.5	15.5	22	0.50	0.25	0.25	0.66	0.33	0
6	62	32.4	22.5	31	30	0.5	0.2	0.3	0.5	0	0.5

For the calculation of $C_{hlg}(Y)$ the following relations can be applied:

$$C_{hlg}(Y) = V_{hlg}(Y) \cdot 2p_{h.} - A_{hlg}(Y) \cdot \frac{n_{h.}}{N} \tag{48}$$

where

$$\begin{cases} V_{hlg}(Y) & = \frac{M_g(Y) - \bar{M}_{hl}(Y)}{M(Y)} \cdot \frac{n_{.g}}{N} \cdot p(l|h) \\ A_{hlg}(Y) & = \frac{M_g(Y) - M_{hl}(Y)}{M(Y)} \cdot \frac{n_{.g}}{N} \cdot f(l|h). \end{cases} \tag{49}$$

Table 10 shows the calculation for $C_{1lg}(Y) = V_{1lg}(Y) \cdot 2p_{1.} - A_{1lg}(Y) \cdot \frac{n_{1.}}{N}$. The values $V_{1lg}(Y)$ and $A_{1lg}(Y)$ are those reported in Zenga (2016a), Tables 7 and 8, while $2p_{1.} = 0,4$ and $\frac{n_{1.}}{N} = 0.2$.

TABLE 10. - *Calculation of $C_{1lg}(Y) = V_{1lg}(Y) \cdot 0.4 - A_{1lg}(Y) \cdot 0.2$.*

$2 \cdot p_{1.} = 0.4$		l		
$\frac{n_{1.}}{N} = 0.2$		1	2	3
g	1	$0.2533 \cdot 0.4 - 0.2533 \cdot 0.2 =$ 0.05066	$0 - 0 =$	$0.2533 \cdot 0.4 - 0.2533 \cdot 0.2 =$ 0.05066
	2	$0.0683 \cdot 0.4 - 0.0683 \cdot 0.2 =$ 0.01366	$0 - 0 = 0$	$0.0683 \cdot 0.4 - 0.0683 \cdot 0.2 =$ 0.01366
	3	$0.145 \cdot 0.4 - 0.145 \cdot 0.2 =$ 0.02900	$0 - 0 = 0$	$0.145 \cdot 0.4 - 0.145 \cdot 0.2 =$ 0.02900
				$G_1(Y) = 0.1866$

5. JOINT DECOMPOSITION BY SOURCES AND SUBPOPULATIONS OF $G_h(Y)$ AND $G(Y)$

To get the joint decomposition we need, for each subpopulation g: the cell means $M_{hg}(X_j)$ and $M_{hg}(Y)$, the lower means $\bar{M}_{hg}(X_j)$ and $\bar{M}_{hg}(Y)$, and the subpopulation means $M_g(X_j)$ and $M_g(Y)$.

5.1 *Means and lower means of the sources for subpopulations*

In the case $n_{hg} > 0$, let us denote with $x_{hgjd}, \quad d = 1, \ldots, n_{hg}$, the values of X_j observable on each of the n_{hg} units of the subpopulation g with total income $Y = y_h$. Obviously, in this case $(n_{hg} > 0)$, $\sum_{j=1}^{c} x_{hgjd} = y_h, \forall d = 1, \ldots, n_{hg}$. The sums of the values of X_j and of Y, observable on each of the n_{hg} units with $Y = y_h$, are denoted by $S_{hg}(X_j)$ and $S_{hg}(Y)$ respectively; note that

$$S_{hg}(X_j) = \sum_{d=1}^{n_{hg}} x_{hgjd} \tag{50}$$

and

$$S_{hg}(Y) = \sum_{d=1}^{n_{hg}} y_h = y_h \cdot n_{hg}. \tag{51}$$

Moreover,

$$\begin{aligned} \sum_{j=1}^{c} S_{hg}(X_j) &= \sum_{j=1}^{c} \sum_{d=1}^{n_h g} x_{hgjd} = \sum_{d=1}^{n_h g} \sum_{j=1}^{c} x_{hgjd} \\ &= \sum_{d=1}^{n_h g} y_h = S_{hg}(Y) = y_h \cdot n_{hg}. \end{aligned} \tag{52}$$

In the case $n_{hg} > 0$,

$$M_{hg}(X_j) = \frac{S_{hg}(X_j)}{n_{hg}} \tag{53}$$

and

$$M_{hg}(Y) = \frac{S_{hg}(Y)}{n_{hg}} = y_h, \tag{54}$$

are the corresponding means of X_j and Y. From the last two relations we have

$$\sum_{j=1}^{c} M_{hg}(X_j) = M_{hg}(Y) = y_h. \tag{55}$$

In the case $n_{hg} = 0$, we set

$$S_{hg}(X_j) = S_{hg}(Y) = M_{hg}(X_j) = M_{hg}(Y) = 0. \tag{56}$$

With these two assumptions the relations $\sum_{j=1}^{c} S_{hg}(X_j) = S_{hg}(Y)$ and $\sum_{j=1}^{c} M_{hg}(X_j) = M_{hg}(Y)$ are extended to the case $n_{hg} = 0$. Let

$$Q_{hg}(X_j) = \sum_{t=1}^{h} S_{tg}(X_j), \tag{57}$$

be the sum of the values of X_j observable on each of the P_{hg} units of the subpopulation g with $Y \leq y_h$. Moreover: $T_g(X_j) = Q_{rg}(X_j) = \sum_{h=1}^{r} S_{hg}(X_j)$ is the sum of the $n_{.g}$ values of X_j of the subpopulation g, and

$$M_g(X_j) = \frac{T_g(X_j)}{n_{.g}}. \tag{58}$$

is the corresponding mean. Now from (57)

$$\sum_{j=1}^{c} Q_{hg}(X_j) = \sum_{j=1}^{c}\sum_{t=1}^{h} S_{tg}(X_j) = \sum_{t=1}^{h}\sum_{j=1}^{c} S_{tg}(X_j) =$$
$$= \sum_{t=1}^{h} S_{tg}(Y) = Q_{hg}(Y). \tag{59}$$

In coherence with (35) and (36), we define the lower means $\bar{M}_{hg}(X_j)$ as follows

$$\bar{M}_{hg}(X_j) = \begin{cases} M_{o(g)g}(X_j), & \text{for } h < o(g) \\ Q_{hg}(X_j)/P_{hg}, & \text{for } h \geq o(g) \end{cases} \tag{60}$$

From (35) we know that $o(g) = \min h : \ n_{hg} > 0$. Consequently, for $h < o(g)$,

$$\bar{M}_{hg}(X_j) = M_{o(g)g}(X_j) = \frac{S_{o(g)g}(X_j)}{n_{o(g)g}}.$$

Finally,

$$\sum_{j=1}^{c} \bar{M}_{hg}(X_j) = \begin{cases} \sum_{j=1}^{c} M_{o(g)g}(X_j) = M_{o(g)g}(Y), & \text{for } h < o(g) \\ \sum_{j=1}^{c} \frac{Q_{hg}(X_j)}{P_{hg}} = \frac{Q_{hg}(Y)}{P_{hg}} = \bar{M}_{hg}(Y) & \text{for } h \geq o(g) \end{cases}$$
$$= \bar{M}_{hg}(Y), h = 1, \ldots, r \tag{61}$$

and

$$M_g(Y) = \sum_{j=1}^{c} M_g(X_j). \tag{62}$$

For the example at hand, Table 11 reports the sums $S_{hg}(X_j)$ and $S_{hg}(Y) = \sum_{j=1}^{c} S_{hg}(X_j)$. Table 12 reports the means: $M_{hg}(X_j)$, $M_{hg}(Y)$; $M_{h.}(X_j)$, $M_{h.}(Y)$; $M_g(X_j)$,$M_g(Y)$; $M(X_j)$, $M(Y)$. Table 13 reports the cumulative incomes $Q_{hg}(X_j)$ and $Q_{h.}(X_j)$ of the sources X_j, and the cumulative incomes $Q_{hg}(Y)$ and $Q_{h.}(Y)$ of the total Y. Finally, Table 14 reports the lower means: $\bar{M}_{hg}(X_j)$ and $\bar{M}_{hg}(Y)$, and $\bar{M}_{h.}(X_j)$ and $\bar{M}_{h.}(Y)$.

TABLE 11. - *Sums:* $S_{hg}(X_j), S_{hg}(Y); S_{h.}(X_j), S_{h.}(Y); T_g(X_j), T_g(Y); T(X_j), T(Y)$

Subpop. g		h = 1	2	3	4	5	6	$T_g(X_j), T_g(Y)$
1	$S_{h1}(X_1)$	1.5	0	5	0	18	10	$34.5 = T_1(X_1)$
	$S_{h1}(X_2)$	0	0	9	0	24	30	$63 = T_1(X_2)$
	$S_{h1}(X_3)$	0.5	0	10	0	32	22	$64.5 = T_1(X_3)$
	$S_{h1}(Y)$	2	0	24	0	74	62	$162 = T_1(Y)$
2	$S_{h2}(X_1)$	0	1	0	0	12	0	$13 = T_2(X_1)$
	$S_{h2}(X_2)$	0	3	0	0	15	0	$18 = T_2(X_2)$
	$S_{h2}(X_3)$	0	4	0	0	10	0	$14 = T_2(X_3)$
	$S_{h2}(Y)$	0	8	0	0	37	0	$45 = T_2(Y)$
3	$S_{h3}(X_1)$	0.5	0	0	4	0	18	$22.5 = T_3(X_1)$
	$S_{h3}(X_2)$	0.0	0	0	15	0	20	$35 = T_3(X_2)$
	$S_{h3}(X_3)$	1.5	0	0	10	0	24	$35.5 = T_3(X_3)$
	$S_{h3}(Y)$	2	0	0	29	0	62	$93 = T_3(Y)$
Whole pop.	$S_{h.}(X_1)$	2	1	5	4	30	28	$70 = T(X_1)$
	$S_{h.}(X_2)$	0	3	9	15	39	50	$116 = T(X_2)$
	$S_{h.}(X_3)$	2	4	10	10	42	46	$114 = T(X_3)$
	$S_{h.}(Y)$	4	8	24	29	111	124	$300 = T(Y)$

TABLE 12. - *Means:* $M_{hg}(X_j),\ M_{hg}(Y);\ M_{h.}(X_j),\ M_{h.}(Y);\ M_g(X_j),\ M_g(Y);$ $M(X_j),\ M(Y)$

Subpop. g		h = 1	2	3	4	5	6	$M_g(X_j), M_g(Y)$
1	$M_{h1}(X_1)$	1.5	0	5	0	9	10	$6.9 = M_1(X_1)$
	$M_{h1}(X_2)$	0	0	9	0	12	30	$12.6 = M_1(X_2)$
	$M_{h1}(X_3)$	0.5	0	10	0	16	22	$12.9 = M_1(X_3)$
	$M_{h1}(Y)$	2	0	24	0	37	62	$32.4 = M_1(Y)$
2	$M_{h2}(X_1)$	0	1	0	0	12	0	$6.5 = M_2(X_1)$
	$M_{h2}(X_2)$	0	3	0	0	15	0	$9 = M_2(X_2)$
	$M_{h2}(X_3)$	0	4	0	0	10	0	$7 = M_2(X_3)$
	$M_{h2}(Y)$	0	8	0	0	37	0	$22.5 = M_2(Y)$
3	$M_{h3}(X_1)*$	0.5	0	0	4	0	18	$7.5 = M_3(X_1)$
	$M_{h3}(X_2)$	0	0	0	15	0	20	$11.66 = M_3(X_2)$
	$M_{h3}(X_3)$	1.5	0	0	10	0	24	$11.83 = M_3(X_3)$
	$M_{h3}(Y)$	2	0	0	29	0	62	$31 = M_3(Y)$
Whole pop.	$M_{h.}(X_1)$	1	1	5	4	10	14	$7 = M(X_1)$
	$M_{h.}(X_2)$	0	3	9	15	13	10	$11.6 = M(X_2)$
	$M_{h.}(X_3)$	1	4	10	10	14	23	$11.4 = M(X_3)$
	$M_{h.}(Y)$	2	8	24	29	37	62	$30 = M(Y)$

TABLE 13. - *Cumulative incomes $Q_{hg}(X_j)$ and $Q_{h.}(X_j)$ of the sources X_j, and cumulative incomes $Q_{hg}(Y)$ and $Q_{h.}(Y)$ of the total Y*

Subpop. g		h: 1	2	3	4	5	6
1	$Q_{h1}(X_1)$	1.5	1.5	6.5	6.5	24.5	34.5
	$Q_{h1}(X_2)$	0.0	0.0	9.0	9.0	33.0	63.0
	$Q_{h1}(X_3)$	0.5	0.5	10.5	10.5	42.5	64.5
	$Q_{h1}(Y)$	2	2	26	26	100	162
2	$Q_{h2}(X_1)$	0	1	1	1	13	13
	$Q_{h2}(X_2)$	0	3	3	3	18	18
	$Q_{h2}(X_3)$	0	4	4	4	14	14
	$Q_{h2}(Y)$	0	8	8	8	45	45
3	$Q_{h3}(X_1)$	0.5	0.5	0.5	4.5	4.5	22.5
	$Q_{h3}(X_2)$	0.0	0.0	0.0	15.0	15.0	35.0
	$Q_{h3}(X_3)$	1.5	1.5	1.5	11.5	11.5	35.5
	$Q_{h3}(Y)$	2	2	2	31	31	93
Whole pop.	$Q_{h.}(X_1)$	2	3	8	12	42	70
	$Q_{h.}(X_2)$	0	3	12	27	66	116
	$Q_{h.}(X_3)$	2	6	16	26	68	114
	$Q_{h.}(Y)$	4	12	36	65	176	300

TABLE 14. - *Lower means: $\bar{M}_{hg}(X_j)$ and $\bar{M}_{hg}(Y)$, and $\bar{M}_{h.}(X_j)$ and $\bar{M}_{h.}(Y)$*

g	o(g)		h: 1	2	3	4	5	6
1	1	$\bar{M}_{h1}(X_1)$	1.5	1.5	3.25	3.25	6.125	6.9
		$\bar{M}_{h1}(X_2)$	0.0	0.0	4.50	4.50	8.250	12.6
		$\bar{M}_{h1}(X_3)$	0.5	0.5	5.25	5.25	10.625	12.9
		$\bar{M}_{h1}(Y)$	2	2	13	13	2.5	32.4
2	2	$\bar{M}_{h2}(X_1)$	1	1	1	1	6.5	6.5
		$\bar{M}_{h2}(X_2)$	3	3	3	3	9.0	9.0
		$\bar{M}_{h2}(X_3)$	4	4	4	4	7.0	7.0
		$\bar{M}_{h2}(Y)*$	8	8	8	8	22.5	22.5
3	1	$\bar{M}_{h3}(X_1)$	0.5	0.5	0.5	2.25	2.25	7.50
		$\bar{M}_{h3}(X_2)$	0.0	0.0	0.0	7.50	7.50	11.66
		$\bar{M}_{h3}(X_3)$	1.5	1.5	1.5	5.75	5.75	11.83
		$\bar{M}_{h3}(Y)$	2	2	2	15.5	15.5	31
Whole pop		$\bar{M}_{h.}(X_1)$	1	1	2	2.4	5.25	7.0
		$\bar{M}_{h.}(X_2)$	0	1	3	5.4	8.25	11.6
		$\bar{M}_{h.}(X_3)$	1	2	4	5.2	8.50	11.4
		$\bar{M}_{h.}(Y)$	2	4	9	13	22	30

5.2 *From the $k \times k$ decomposition by subpopulations of the Gini measures to the corresponding $k \times k \times c$ joint decomposition by subpopulations and sources*

By the use of the following relations (62), (61),and (55)

$$M_g(Y) = \sum_{j=1}^{c} M_g(X_j) \tag{62}$$

$$\bar{M}_{hl}(Y) = \sum_{j=1}^{c} \bar{M}_{hl}(X_j) \tag{61}$$

$$y_h = M_{hl}(Y) = \sum_{j=1}^{c} M_{hl}(X_j) \tag{55}$$

in the decomposition by subpopulations (44) of $G_h(Y)$, the basic $k \times k \times c$ joint decomposition by subpopulations and sources of $G_h(Y)$ is obtained

$$G_h(Y) = \sum_{l=1}^{k} \sum_{g=1}^{k} \sum_{j=1}^{c} C_{hlg}(X_j) \tag{63}$$

where

$$C_{hlg}(X_j) = \frac{M_g(X_j) - \bar{M}_{hl}(X_j)}{M(Y)} \frac{n_{.g}}{N} p(l|h) 2p_{h.} - \frac{M_g(X_j) - M_{hl}(X_j)}{M(Y)} \frac{n_{.g}}{N} f(l|h) \frac{n_{h.}}{N}. \tag{64}$$

For the calculation of $C_{hlg}(X_j)$ the following relations are useful:

$$C_{hlg}(X_j) = V_{hlg}(X_j) \cdot 2p_{h.} - A_{hlg}(X_j) \cdot \frac{n_{h.}}{N} \tag{65}$$

where

$$\begin{cases} V_{hlg}(X_j) & = \frac{M_g(X_j) - \bar{M}_{hl}(X_j)}{M(Y)} \cdot \frac{n_{.g}}{N} \cdot p(l|h) \\ A_{hlg}(X_j) & = \frac{M_g(X_j) - M_{hl}(X_j)}{M(Y)} \cdot \frac{n_{.g}}{N} \cdot f(l|h). \end{cases} \tag{66}$$

Tables 15, 16 and 17 show the calculations for $V_{1lg}(X_1)$, $V_{1lg}(X_2)$ and $V_{1lg}(X_3)$. Note that $\sum_{l=1}^{k} \sum_{g=1}^{k} V_{1lg}(X_j) = V_1(X_j)$, and that $\sum_{j=1}^{3} V_{1lg}(X_j) = V_{1lg}(Y)$. Tables 18, 19 and 20 report the calculations for $A_{1lg}(X_j)$, $j = 1, 2, 3$. Note that $\sum_{l=1}^{k} \sum_{g=1}^{k} A_{1lg}(X_j) = A_1(X_j)$, and that $\sum_{j=1}^{3} A_{1lg}(X_j) = A_{1lg}(Y)$. Finally, Tables 21, 22 and 23 report the values of $C_{1lg}(X_j)$ and of $C_1(X_j)$.

TABLE 15. - *Calculation of* $V_{1lg}(X_1) = \frac{M_g(X_1)-\bar{M}_{1l}(X_1)}{M(Y)} \cdot \frac{n_{.g}}{N} \cdot p(l|1)$

$h = 1$ $M(Y) = 30$			l			$\frac{n_{.g}}{N}$
			1	2	3	
			$\bar{M}_{11}(X_1) = 1.5$	$\bar{M}_{12}(X_1) = 1$	$\bar{M}_{13}(X_1) = 0.5$	
g	1	$M_1(X_1) = 6.9$	$0.18 \cdot 0.25 = 0.0450$	0.0	$0.2133 \cdot 0.25 = 0.0533$	0.5
	2	$M_2(X_1) = 6.5$	$0.166 \cdot 0.1 = 0.0166$	0.0	$0.2 \cdot 0.1 = 0.0200$	0.2
	3	$M_3(X_1) = 7.5$	$0.2 \cdot 0.15 = 0.0300$	0.0	$0.233 \cdot 0.15 = 0.0350$	0.3
$p(l\|1)$			0.5	0.0	0.5	1.0

TABLE 16. - *Calculation of* $V_{1lg}(X_2) = \frac{M_g(X_2)-\bar{M}_{1l}(X_2)}{M(Y)} \cdot \frac{n_{.g}}{N} \cdot p(l|1)$

$h = 1$ $M(Y) = 30$			l			$\frac{n_{.g}}{N}$
			1	2	3	
			$\bar{M}_{11}(X_2) = 0$	$\bar{M}_{12}(X_2) = 3$	$\bar{M}_{13}(X_2) = 0$	
g	1	$M_1(X_2) = 12.6$	$0.42 \cdot 0.25 = 0.1050$	0.0	$0.42 \cdot 0.25 = 0.1050$	0.5
	2	$M_2(X_2) = 9$	$0.3 \cdot 0.1 = 0.0300$	0.0	$0.3 \cdot 0.1 = 0.0300$	0.2
	3	$M_3(X_2) = 11.666$	$0.388 \cdot 0.15 = 0.0583$	0.0	$0.388 \cdot 0.15 = 0.0583$	0.3
$p(l\|1)$			0.5	0.0	0.5	1.0

TABLE 17. - *Calculation of* $V_{1lg}(X_3) = \frac{M_g(X_3)-\bar{M}_{1l}(X_3)}{M(Y)} \cdot \frac{n_{.g}}{N} \cdot p(l|1)$

$h = 1$ $M(Y) = 30$			l			$\frac{n_{.g}}{N}$
			1	2	3	
			$\bar{M}_{11}(X_3) = 0.5$	$\bar{M}_{12}(X_3) = 4$	$\bar{M}_{13}(X_3) = 1.5$	
g	1	$M_1(X_3) = 12.9$	$0.4133 \cdot 0.25 = 0.10330$	0.0	$0.38 \cdot 0.25 = 0.09500$	0.5
	2	$M_2(X_3) = 7$	$0.2166 \cdot 0.1 = 0.02166$	0.0	$0.1833 \cdot 0.1 = 0.01833$	0.2
	3	$M_3(X_3) = 11.83$	$0.3777 \cdot 0.15 = 0.05660$	0.0	$0.3444 \cdot 0.15 = 0.05166$	0.3
$p(l\|1)$			0.5	0.0	0.5	1.0

TABLE 18. - *Calculation of* $A_{1lg}(X_1) = \frac{M_g(X_1) - M_{1l}(X_1)}{M(Y)} \cdot \frac{n_{.g}}{N} \cdot f(l|1)$

$h = 1$, $M(Y) = 30$			1			$\frac{n_{.g}}{N}$
			1	2	3	
			$M_{11}(X_1) = 1.5$	$M_{12}(X_1) = 0$	$M_{13}(X_1) = 0.5$	
g	1	$M_1(X_1) = 6.9$	$0.18 \cdot 0.25 = 0.0450$	0.0	$0.213 \cdot 0.25 = 0.0533$	0.5
	2	$M_2(X_1) = 6.5$	$0.166 \cdot 0.1 = 0.0166$	0.0	$0.2 \cdot 0.1 = 0.0200$	0.2
	3	$M_3(X_1) = 7.5$	$0.2 \cdot 0.15 = 0.0300$	0.0	$0.233 \cdot 0.15 = 0.1450$	0.3
$f(l|1)$			0.5	0.0	0.5	1.0

TABEL 19. - *Calculation of* $A_{1lg}(X_2) = \frac{M_g(X_2) - M_{1l}(X_2)}{M(Y)} \cdot \frac{n_{.g}}{N} \cdot f(l|1)$

$h = 1$, $M(Y) = 30$			l			$\frac{n_{.g}}{N}$
			1	2	3	
			$M_{11}(X_2) = 0$	$M_{12}(X_2) = 0$	$M_{13}(X_2) = 0$	
g	1	$M_1(X_2) = 12.6$	$0.42 \cdot 0.25 = 0.1050$	0.0	$0.42 \cdot 0.25 = 0.105$	0.5
	2	$M_2(X_2) = 9$	$0.3 \cdot 0.1 = 0.0300$	0.0	$0.3 \cdot 0.1 = 0.03$	0.2
	3	$M_3(X_2) = 11.666$	$0.388 \cdot 0.15 = 0.0583$	0.0	$0.388 \cdot 0.15 = 0.0583$	0.3
$f(l|1)$			0.5	0.0	0.5	1.0

TABLE 20. - *Calculation of* $A_{1lg}(X_3) = \frac{M_g(X_3) - M_{1l}(X_3)}{M(Y)} \cdot \frac{n_{.g}}{N} \cdot f(l|1)$

$h = 1$, $M(Y) = 30$			1			$\frac{n_{.g}}{N}$
			1	2	3	
			$M_{11}(X_3) = 0.5$	$M_{11}(X_3) = 0$	$M_{11}(X_3) = 1.5$	
g	1	$M_1(X_3) = 12.9$	$0.4133 \cdot 0.25 = 0.1033$	0.0	$0.38 \cdot 0.25 = 0.09500$	0.5
	2	$M_2(X_3) = 7$	$0.2166 \cdot 0.1 = 0.0216$	0.0	$0.1833 \cdot 0.1 = 0.01833$	0.2
	3	$M_3(X_3) = 11.83$	$0.377 \cdot 0.15 = 0.0566$	0.0	$0.3444 \cdot 0.15 = 0.05166$	0.3
$f(l|1)$			0.5	0.0	0.5	1.0

TABLE 21. - *Values of* $C_{1lg}(X_1) = V_{1lg}(X_1) \cdot 0.4 - A_{1lg}(X_1) \cdot 0.2$ *and of* $C_1(X_1)$

$2p_{1.} = 0.4$		l		
$\frac{n_{1.}}{N} = 0.2$		1	2	3
g	1	0.0090	0	0.0106
	2	0.0033	0	0.0040
	3	0.0060	0	0.0020

$0.04 = C_1(X_1)$

TABLE 22. - *Values of* $C_{1lg}(X_2) = V_{1lg}(X_2) \cdot 0.4 - A_{1lg}(X_2) \cdot 0.2$ *and of* $C_1(X_2)$

$2p_{1.} = 0.4$		l		
$\frac{n_{1.}}{N} = 0.2$		1	2	3
g	1	0.0210	0	0.0210
	2	0.0060	0	0.0060
	3	0.0116	0	0.0116

$0.0773 = C_1(X_2)$

TABLE 23. - *Values of* $C_{1lg}(X_3) = V_{1lg}(X_3) \cdot 0.4 - A_{1lg}(X_3) \cdot 0.2$ *and of* $C_1(X_3)$

$2p_{1.} = 0.4$		l		
$\frac{n_{1.}}{N} = 0.2$		1	2	3
g	1	0.0206	0	0.0190
	2	0.0043	0	0.0036
	3	0.0113	0	0.0103

$0.0693 = C_1(X_3)$

6. DECOMPOSITION BY SOURCES FOR EACH SUBPOPULATION

The point inequality index of the subpopulation l is given by:

$$G_{hl}(Y) = \frac{M_l(Y) - \bar{M}_{hl}(Y)}{M_l(Y)} \cdot 2 \cdot p_{hl} - \frac{M_l(Y) - y_h}{M_l(Y)} \cdot \frac{n_{hl}}{n_{.l}}, \tag{67}$$

where $p_{hl} = \frac{P_{hl}}{n_{.l}}$. The synthetic Gini inequality index of the subpopulation l is given by:

$$G_{.l}(Y) = \sum_{h=1}^{r} G_{hl}(Y) \cdot \frac{n_{hl}}{n_{.l}}. \tag{68}$$

It can be useful to decompose by sources the point $G_{hl}(Y)$ and the synthetic $G_{.l}(Y)$ inequality indexes. By the use of (62), (61) and (54) in the numerator of (67) we have:

$$G_{hl}(Y) = \sum_{j=1}^{c} C_{hl}(X_j), \tag{69}$$

where

$$C_{hl}(X_j) = \frac{M_l(X_j) - \bar{M}_{hl}(X_j)}{M_l(Y)} 2 \cdot p_{hl} - \frac{M_l(X_j) - M_{hl}(X_j)}{M_l(Y)} \cdot \frac{n_{hl}}{n_{.l}} \tag{70}$$

is the contribution of X_j to $G_{hl}(Y)$. Finally, putting the decomposition (69) in (68), the decomposition by sources of $G_{.l}(Y)$ is obtained:

$$G_{.l}(Y) = \sum_{h=1}^{r} \left\{ \sum_{j=1}^{c} C_{hl}(X_j) \right\} \cdot \frac{n_{hl}}{n_{.l}} = \sum_{j=1}^{c} C_{.l}(X_j) \tag{71}$$

where

$$C_{.l}(X_j) = \sum_{h=1}^{r} C_{hl}(X_j) \cdot \frac{n_{hl}}{n_{.l}}. \tag{72}$$

7. APPLICATION

The data used in this application are supplied by the 2014 Central Bank of Italy sample survey on household income and wealth (Bank of Italy 2016). This survey covers $N = 8156$ households. In this paper we deal with the household net disposable income Y, that is the sum of: the payroll income X_1, the pensions and net transfers X_2, the net self employment income X_3, and the property incomes X_4. Table (24) reports the frequency distribution of the $N = 8156$ households according to their number of persons D. This Table shows that there are only 452 households with $D \geq 5$, consequency we have aggregated all the 1674 households with $D \geq 4$. Thus the 8156 households have been partitioned into the following $k = 4$ subpopulations: households with $D = 1$; households with $D = 2$; households with $D = 3$; households with $D \geq 4$. In all computations that follow we consider the weights $w_i > 0$ $(i = 1, 2, \ldots, 8156; W = \sum w_i = 8156)$, supplied by the Central Bank of Italy for each household; these weights are based on the inverse of household's probability of inclusion in the sample (For further details see Bank of Italy 2016). Now we remark that, in the following sections we will not use the notation related to the weights w_i, but for simplicity's sake we will continue the use of the notation of the previous sections. Thus, to denote the sum of the weights of the n_{hl} households with total income $Y = y_h$ of the subpopulation l we will use n_{hl} instead of w_{hl}: $l = 1$ for $D = 1$, $l = 2$ for $D = 2$, $l = 3$ for $D = 3$, $l = 4$ for $D \geq 4$. Note that the frequency distribution of the total income Y has $r = 7400$ different values, 7117 have frequency $n_{h.} = 1$. For an application on EU income data see Pasquazzi and Zenga (2018).

TABLE 24. - *Frequency distribution of the* $N = 8156$ *households according to their number of persons D*

D	1	2	3	4	5	6	7	8	9	10	Tot
frequency	2394	2588	1500	1222	340	85	21	4	1	1	8156

7.1 *Aggegate characteristics in four Italian household groups sorted by their number of persons*

Table 25 reports for each subpopulations: the arithmetic means $M_l(X_j)$ and $M_l(Y)$ of the components X_j and of the total income Y; the shares $\gamma_{.l}(X_j) = M_l(X_j)/M_l(Y)$; the shares $\gamma_{..}(X_j) = M(X_j)/M(Y)$; the synthetic index $G_{.l}(Y)$; the sum of the weights $n_{.l}$; the relative weights $n_{.l}/N$; the shares $T_l(Y)/T(Y)$. Table 25 shows that: the mean value of the total income Y increases with the nunber of persons D; however the difference between the means of the households with $D = 3$ and $D \geq 4$ is negligible. Now it is useful to remember that the shares $\gamma_{.l}(X_j)$ of the four subpopulations are related to the corresponding shares$\gamma_{..}(X_j)$ of the whole country by the relation: $\gamma_{..}(X_j) = \sum_{l=1}^{k} \gamma_{.l}(X_j) \cdot T_l(Y)/T(Y)$. Thus, Table 25 shows that: for the payroll income X_1 the shares of the households with $(D = 3)$ and $(D \geq 4)$ are a lot greater than those of the househols with $(D = 1)$ and $(D = 2)$; for the self employment income X_3 the shares of the households with $(D = 3)$ and $(D \geq 4)$ are greater than those of the househols with $(D = 1)$ and $(D = 2)$; for the pensions and net transfers income X_2 the shares of the households with $(D = 3)$ and $(D \geq 4)$ are a lot lower than those of the househols with $(D = 1)$ and $(D = 2)$; for the property income X_4 the shares of the households with $(D = 3)$ and $(D \geq 4)$ are lower than those of the househols with $(D = 1)$ and $(D = 2)$. Moreover, Table 25 shows that the values of the synthetic Gini index $G_{.l}(Y)$ of the households with $(D = 3)$ and $(D \geq 4)$ are a little bit greater than those of the househols with $(D = 1)$ and $(D = 2)$; in addition the value of the synthetic Gini index $G(Y)$ of the whole country is greater than the corresponding index of the households with $(D \geq 4)$. The synthetic inequality index $G(Y) = 0.35$ means that in the whole population, on (weighted) average, the lower mean is equal to the (1-0.35) $\cdot 100 = 65\%$ of the mean $M(Y) = 30525.2$.

Figure 3 displays the graphs of the point inequality measures for: *a*) the whole population; *b*) the four subpopulations $l = 1$ for $D = 1$, $l = 2$ for $D = 2$, $l = 3$ for $D = 3$ and $l = 4$ for $D \geq 4$. For the subpopulation l the abscissas and the ordinates are

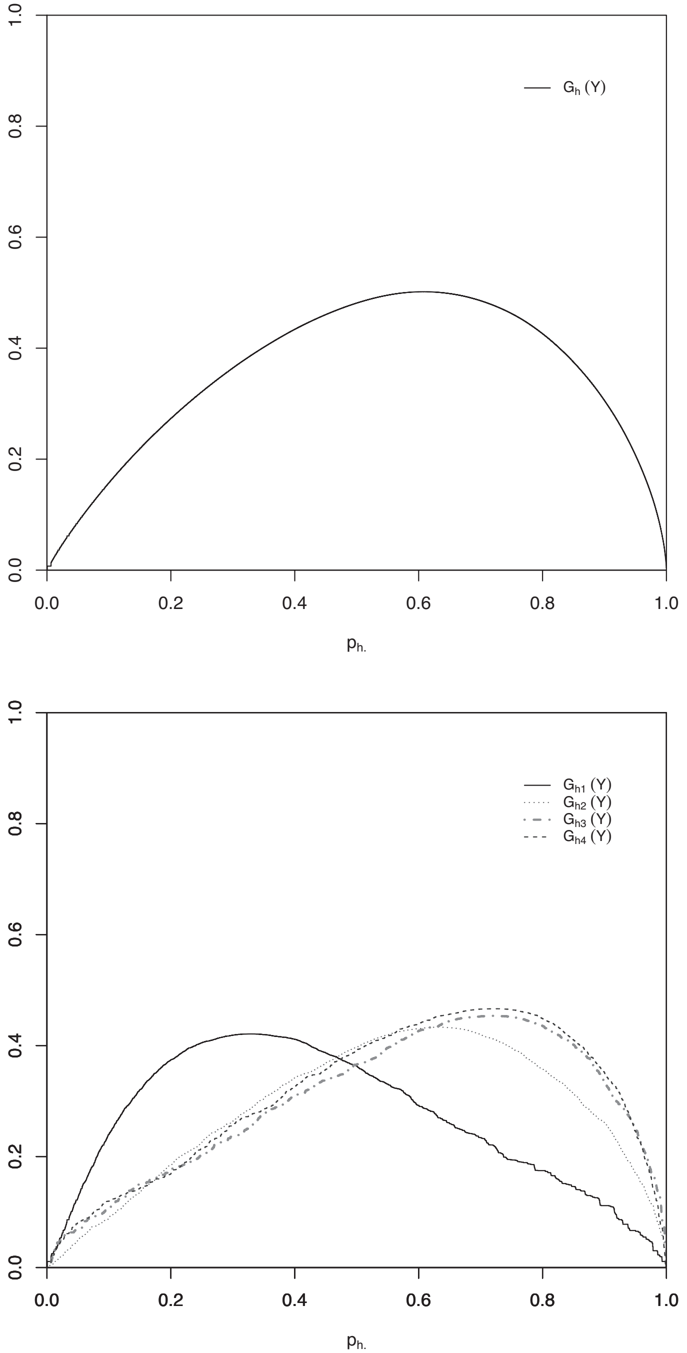

FIGURE 3. - *Graphs of the point measures for households with different number of persons*

$$p_{hl} = \frac{P_{hl}}{n_{.l}} \quad \text{and} \quad G_{hl}(Y), \forall h = 1, \ldots, r,$$

while for the whole population the abscissas and the ordinates are

$$p_{h.} = \frac{P_{h.}}{N} \quad \text{and} \quad G_h(Y), \forall h = 1, \ldots, r\,.$$

TABLE 25. - *Some aggregate characteristics for households with different number of persons D*

	$D=1$ $l=1$	$D=2$ $l=2$	$D=3$ $l=3$	$D\geq 4$ $l=4$	Italy
$n_{.l}$	2392.3	2226.2	1578.03	1959.46	8156
$n_{.l}/N$	0.2933	0.273	0.1935	0.2402	1.0
$G_{.l}(Y)$	0.3055	0.3075	0.3254	0.3271	$0.35 = G(Y)$
Means of Y and X_j, and shares $T_l(Y)/T(Y)$					
X_1	5433.8	8034.3	17261.7	21620.9	12321.0
X_2	7137.0	14029.9	7867.5.	3913.5	8385.3
X_3	1549.1	2633.3	5334.2	5443.3	3513.0
X_4	5087.2	7197.7	6808.0	6381.1	6305.9
Y	19207.1	31895.2	37265.4	37358.8	30525.2
$T_l(Y)/T(Y)$	0.1846	0.2852	0.2362	0.294	1.0
Shares: $\gamma_{.l}(X_j)$ and $\gamma_{..}(X_j)$					
$\gamma_{.l}(X_1)$	0.2829	0.2519	0.4632	0.5787	$0.4036 = \gamma_{..}(X_1)$
$\gamma_{.l}(X_2)$	0.3716	0.4399	0.2111	0.1048	$0.2747 = \gamma_{..}(X_2)$
$\gamma_{.l}(X_3)$	0.0807	0.0826	0.1431	0.1457	$0.1151 = \gamma_{..}(X_3)$
$\gamma_{.l}(X_4)$	0.2649	0.2257	0.1825	0.1708	$0.2066 = \gamma_{..}(X_4)$
	1.0000	1.0000	1.0000	1.0000	1.0000

7.2 *Joint decomposition by household number of persons and sources of the point and the synthetic Gini index*

The decomposition of $G_h(Y)$ and $G(Y)$ by subpopulations-Geographical Areas- is widely illustrated in Zenga (2016a), while the case of the decomposition by sources is presented in Zenga (2013). In this section we illustrate the decompositions of the point measure $G_{h(p)}(Y) = G_{(p)}(Y)$ for three values of p, and the decompositions of the synthetic index $G(Y) = 0.35$. For p we have chosen the following values: $p = 0.10$, because $G_{(0.10)}(Y) = 0.1568$ "compares" the mean income of the poorest 10% households with the mean $M(Y)$ of the whole population; $p = 0.50$, because $G_{(0.50)}(Y) = 0.4813$ "compares" the mean income of the households with $Y \leq Median(Y)$ with the mean $M(Y)$; $p = 0.95$, because $G_{(0.95)}(Y) = 0.2072$ "compares" the mean income of the lower group that is the 95% of the whole population with the mean $M(Y)$. Note that: $h(p) = \left\{min\ h : \frac{P_{h.}}{N} \geq p; h = 1, 2, .., r\right\}$.

Tables 26, 27 and 28 report the joint decompositions of the three mentioned Gini point measures, as well as all the values necessary for their computations, while Table 29 reports the joint decompositions of the synthetic Gini index; it is useful to remember that the contributions reported in this latter table are the weighted arithmetic means of the corresponding contributions of the r point indexes $G_h(Y)$ with weights $n_{h.}/N$.

We remark that for the interpretation of the values repoted in Tables 26, 27 and 28 it is useful to analyze the expression (11) of $G_h(Y)$:

$$G_h(Y) = \frac{M(Y) - \bar{M}_{h.}(Y)}{M(Y)} \cdot 2p_{h.} - \frac{M(Y) - y_h}{M(Y)} \cdot \frac{n_{h.}}{N}$$
$$= V_h(Y) \cdot 2p_{h.} - A_h(Y) \cdot \frac{n_{h.}}{N},$$

where

$$V_h(Y) = \frac{M(Y) - \bar{M}_{h.}(Y)}{M(Y)}$$

is the Bonferroni point measure of inequality: see Zenga and Valli (2016, 2017). Now, we will show that, for the present application, the values of $G_h(Y)$ are fundamentally given by the values of the first term of the right side of $G_h(Y)$. This happens because on average the value of the relative frequency $\frac{n_{h.}}{N}$ is negligible and consequently are negligible the values of the second term of the right side of (11) . In fact, $r = 7400$, $\sum_{h=1}^{r} n_{h.} = N = 8156$ and consequently $M(n_{h.}) = \frac{8156}{7400} = 1.102$ and $M\left(\frac{n_{h.}}{N}\right) = \frac{1.102}{8156} = 0.0001351$. Now we report the details for the calculation of the three Gini point measures.

1. For $p = 0.1$, we have: $h = 440$; $y_h = 10400$; $M(Y) = 30525.2$; $\bar{M}_{h.}(Y) = 6521.2$; $n_{h.} = 7.906$. Thus: $V_h(Y) = 0.7864$; $2p_{h.} = 0.2$; $A_h(Y) = 0.65929$; $\frac{n_{h.}}{N} = 0.00097$.
   Finally, $G_h(Y) = 0.15728 - 0.00064 = 0.15664$.
2. For $p = 0.5$, we have: $h = 3262$; $y_h = 25107.88$; $M(Y) = 30525.2$; $\bar{M}_{h.}(Y) = 15835.87$; $n_{h.} = 1.1932$. Thus: $V_h(Y) = 0.48122$; $2p_{h.} = 1.0$; $A_h(Y) = 0.17747$; $\frac{n_{h.}}{N} = 0.0001463$.
   Finally, $G_h(Y) = 0.48122 - 0.000026 = 0.48119$.
3. $p = 0.95$, we have: $h = 6945$; $y_h = 67436.07$; $M(Y) = 30525.2$; $\bar{M}_{h.}(Y) = 27198.91$; $n_{h.} = 1.1187$. Thus: $V_h(Y) = 0.10897$; $2p_{h.} = 1.9$; $A_h(Y) = -1.209193$; $\frac{n_{h.}}{N} = 0.000137$.
   Finally, $G_h(Y) = 0.207043 + 0.0001658 = 0.207208$.

In conclusion, in the present application, the values of the Gini point index $G_h(Y)$ are very close to the product of the corresponding Bonferroni index $V_h(Y)$ and $2 \cdot p_{h.}$.

TABLE 26. - *Means, lower means, and conditional relative frequencies in the households with different number of persons, and joint decomposition of $G_{h(0.1)}(Y) = 0.1568$ by subpopulations and sources. Calculation based on Bank of Italy 2014 sample survey data*

$p = 0.10; h = 440$	$D = 1$ $l = 1$	$D = 2$ $l = 2$	$D = 3$ $l = 3$	$D \geq 4$ $l = 4$	Italy
$Y \leq 10400$	453.52	125.52	99.37	137.85	$816.26 = P_{h.}$
Total $n_{.l}$	2392.32	2226.18	1578.04	1959.46	8156
n_{hl}	4.3569	2.2548	0.8908	0.4042	$7.9068 = n_{h.}$
Conditional relative frequencies					
$p(l\|h)$	0.5556	0.1538	0.1217	0.1689	1.0000
$f(l\|h)$	0.5510	0.2852	0.1127	0.0511	1.0000
Means and Lower means					
$\bar{M}_{hl}(X_1)$	2130.6	1111.2	1449.3	2094.6	1884.8
$M_l(X_1)$	5433.8	8034.3	17261.7	21620.9	12321.0
$\bar{M}_{hl}(X_2)$	3041.8	3349.5	1986.5	1670.3	2729.0
$M_l(X_2)$	7137.0	14029.9	7867.5.	3913.5	8385.3
$\bar{M}_{hl}(X_3)$	249.3	764.3	180.0	179.3	308.2
$M_l(X_3)$	1549.1	2633.3	5334.2	5443.3	3513.0
$\bar{M}_{hl}(X_4)$	1610.7	1461.7	1665.7	1638.0	1599.1
$M_l(X_4)$	5087.2	7197.7	6802.0	6381.1	6305.9
$\bar{M}_{hl}(Y)$	7032.4	6686.7	5281.5	5582.2	6521.2
$M_l(Y)$	19207.1	31895.2	37265.4	37358.8	30525.2
Joint decomposition					
$C_{h(0.1)lW}(X_1)$	0.0035	0.0019	0.0024	0.0052	$0.0130 = C_{h(0.1).W}(X_1)$
$C_{h(0.1)lB}(X_1)$	0.0334	0.0094	0.006.2	0.0061	$0.0551 = C_{h(0.1).B}(X_1)$
$C_{h(0.1)l.}(X_1)$	0.0369	0.0113	0.0087	0.0113	$0.0682 = C_{h(0.1)..}(X_1)$
$C_{h(0.1)lW}(X_2)$	0.0044	0.0029	0.0009	0.0006	$0.0088 = C_{h(0.1).W}(X_2)$
$C_{h(0.1)lB}(X_2)$	0.0151	0.0021	0.0042	0,0068	$0.0282 = C_{h(0.1).B}(X_2)$
$C_{h(0.1)l.}(X_2)$	0.0195	0.0050	0.0051	0.0074	$0.0370 = C_{h(0.1)..}(X_2)$
$C_{h(0.1)lW}(X_3)$	0.0014	0.0005	0.0008	0.0014	$0.0041 = C_{h(0.1).W}(X_3)$
$C_{h(0.1)lB}(X_3)$	0.0104	0.0022	0.0019	0.0023	$0.0168 = C_{h(0.1).B}(X_3)$
$C_{h(0.1)l.}(X_3)$	0.0118	0.0027	0.0026	0.0037	$0.0208 = C_{h(0.1)..}(X_3)$
$C_{h(0.1)lW}(X_4)$	0.0037	0.0016	0.0008	0.0013	$0.0074 = C_{h(0.1).W}(X_4)$
$C_{h(0.1)lB}(X_4)$	0.0133	0.0033	I0.0029	0.0039	$0.0234 = C_{h(0.1).B}(X_4)$
$C_{h(0.1)l.}(X_4)$	0.0170	0.0048	0.0037	0.0052	$0.0307 = C_{h(0.1)..}(X_4)$
$C_{h(0.1)lW}(Y)$	0.0130	0.0069	0.0049	0.0084	$0.0332 = C_{h(0.1).W}(Y)$
$C_{h(0.1)lB}(Y)$	0.0723	0.0170	0.0152	0.0191	$0.1236 = C_{h(0.1).B}(Y)$
$C_{h(0.1)l.}(Y)$	0.0852	0.0239	0.0201	0.0276	$0.1568 = G_{h(0.1)}(Y)$

TABLE 27. - *Means, lower means and conditional relafive frequencies in the households with different number of persons, and joint decomposition of $G_{h(0.5)}(Y) = 0.4813$ by subpopulations and sources. Calculation based on Bank of Italy 2014 sample survey data*

$p = 0.50; h = 3262$	$D = 1$ $l = 1$	$D = 2$ $l = 2$	$D = 3$ $l = 3$	$D \geq 4$ $l = 4$	Italy
$Y \leq 25107.88$	1926.55	953.04	510.71	688.81	$4079.12 = P_{h.}$
Total $n_{.l}$	2392.32	2226.18	1578.04	1959.46	8156
n_{hl}	1.1932	0	0	0	$1.1932 = n_{h.}$
Conditional relative frequencies					
$p(l\|h)$	0.4723	0.2336	0.1252	0.1689	1.0000
$f(l\|h)$	1	0	0	0	1.0000
Means and Lower means					
$\bar{M}_{hl}(X_1)$	4385.15	4348.58	7964.15	9829.83	5744.11
$M_l(X_1)$	5433.8	8034.3	17261.7	21620.9	12321.0
$\bar{M}_{hl}(X_2)$	6174.45	8670.74	4080.82	2414.06	5860.57
$M_l(X_2)$	7137.0	14029.9	7867.5.	3913.5	8385.3
$\bar{M}_{hl}(X_3)$	772.33	784.06	1295.51	1524.23	967.54
$M_l(X_3)$	1549.1	2633.3	5334.2	5443.3	3513.0
$\bar{M}_{hl}(X_4)$	3546.05	3291.06	2885.53	2716.29	3263.66
$M_l(X_4)$	5087.2	7197.7	6802.0	6381.1	6305.9
$\bar{M}_{hl}(Y)$	14878.00	17094.44	16226.01	16484.4	15835.87
$M_l(Y)$	19207.1	31895.2	37265.4	37358.8	30525.2
Joint decomposition					
$C_{h(0.5)lW}(X_1)$	0.0048	0.0077	0.0074	0.0157	$0.0356 = C_{h(0.5).W}(X_1)$
$C_{h(0.5)lB}(X_1)$	0.1181	0.0533	0.0105	-0.0019	$0.1800 = C_{h(0.5).B}(X_1)$
$C_{h(0.5)l.}(X_1)$	0.1228	0.0610	0.0179	0.0138	$0.2155 = C_{h(0.5)..}(X_1)$
$C_{h(0.5)lW}(X_2)$	0.0044	0.0112	0.0030	0.0020	$0.0206 = C_{h(0.5).W}(X_2)$
$C_{h(0.5)lB}(X_2)$	0.0298	-0.013	0.0147	0.0310	$0.0621 = C_{h(0.5).B}(X_2)$
$C_{h(0.5)l.}(X_2)$	0.0342	-0.002	0.0177	0.0330	$0.0827 = C_{h(0.5)..}(X_2)$
$C_{h(0.5)lW}(X_3)$	0.0035	0.0039	0.0032	0.0052	$0.0158 = C_{h(0.5).W}(X_3)$
$C_{h(0.5)lB}(X_3)$	0.0389	0.0170	0.0059	0.0058	$0.0676 = C_{h(0.5).B}(X_3)$
$C_{h(0.5)l.}(X_3)$	0.0424	0.0209	0.0091	0.0110	$0.0834 = C_{h(0.5)..}(X_3)$
$C_{h(0.5)lW}(X_4)$	0.0070	0.0082	0.0031	0.0049	$0.0232 = C_{h(0.5).W}(X_4)$
$C_{h(0.5)lB}(X_4)$	0.0357	0.0149	0.0109	0.0150	$0.0765 = C_{h(0.5).B}(X_4)$
$C_{h(0.5)l.}(X_4)$	0.0427	0.0231	0.0140	0.0199	$0.0997 = C_{h(0.5)..}(X_4)$
$C_{h(0.5)lW}(Y)$	0.0197	0.0309	0.0167	0.0278	$0.0951 = C_{h(0.5).W}(Y)$
$C_{h(0.5)lB}(Y)$	0.2225	0.0719	0.0420	0.0499	$0.3863 = C_{h(0.5).B}(Y)$
$C_{h(0.5)l.}(Y)$	0.2421	0.1028	0.0587	0.0777	$0.4813 = G_{h(05)}(Y)$

TABLE 28. - *Means, lower means and conditional relafive frequencies in the households with different number of persons, and joint decomposition of $G_{h(0.95)}(Y) = 0.2072$ by subpopulations and sources. Calculation based on Bank of Italy 2014 sample survey data*

$p = 0.95; h = 6945$	$D = 1$ $l = 1$	$D = 2$ $l = 2$	$D = 3$ $l = 3$	$D \geq 4$ $l = 4$	Italy
$Y \leq 67436, 07$	2371.30	2132.52	1457.18	1787.84	$7748, 84 = P_{h.}$
Total $n_{.l}$	2392.32	2226.18	1578.04	1959.46	8156
n_{hl}	0	1.1187	0	0	$1.1187 = n_{h.}$
Conditional relative frequencies					
$p(l\|h)$	0.3060	0.2752	0.1881	0.2307	1.0000
$f(l\|h)$	0	1	0	0	1.0000
Means and Lower means					
$\bar{M}_{hl}(X_1)$	5375.88	7729.25	16018.63	19981.68	11394.82
$M_l(X_1)$	5433.8	8034.3	17261.7	21620.9	12321.0
$\bar{M}_{hl}(X_2)$	7114.79	13301.23	7121.55	3375.20	7955.79
$M_l(X_2)$	7137.0	14029.9	7867.5.	3913.5	8385.3
$\bar{M}_{hl}(X_3)$	1169.75	1683.09	3441.96	3743.84	2332.21
$M_l(X_3)$	1549.1	2633.3	5334.2	5443.3	3513.0
$\bar{M}_{hl}(X_4)$	4900.40	6371.86	5632.23	5217.26	5516.08
$M_l(X_4)$	5087.2	7197.7	6802.0	6381.1	6305.9
$\bar{M}_{hl}(Y)$	18560.81	29085.43	32214.37	32317.97	27198.91
$M_l(Y)$	19207.1	31895.2	37265.4	37358.8	30525.2
Joint decomposition					
$C_{h(0.95)lW}(X_1)$	0.003	0.0014	0.0028	0.0057	$= C_{h(0.95).W}(X_1)$
$C_{h(0.95)lB}(X_1)$	0.1320	0.0772	-0.0461	-0.1157	$= C_{h(0.95).B}(X_1)$
$C_{h(0.95)l.}(X_1)$	0.1323	0.0786	-0.0433	-0.1100	$0.1514 = C_{h(0.95)..}(X_1)$
$C_{h(0.95)lW}(X_2)$	0.0001	0.0034	0.0017	0.0019	$= C_{h(0.95).W}(X_2)$
$C_{h(0.95)lB}(X_2)$	0.0241	-0.0875	0.0131	0.0701	$= C_{h(0.95).B}(X_2)$
$C_{h(0.95)l.}(X_2)$	0.0242	-0.0840	0.0148	0.0720	$0.0572 = C_{h(0.95)..}(X_2)$
$C_{h(0.95)lW}(X_3)$	0.0021	0.0044	0.0043	0.0059	$= C_{h(0.95).W}(X_3)$
$C_{h(0.95)lB}(X_3)$	0.0425	0.0269	-0.0035	-0.0092	$= C_{h(0.95).B}(X_3)$
$C_{h(0.95)l.}(X_3)$	0.0446	0.0313	0.0008	-0.0033	$0.0681 = C_{h(0.95)..}(X_3)$
$C_{h(0.95)lW}(X_4)$	0.0010	0.0039	0.0026	0.0040	$= C_{h(0.95).W}(X_4)$
$C_{h(0.95)lB}(X_4)$	0.0257	-0.0050	0.0052	0.0116	$= C_{h(0.95).B}(X_4)$
$C_{h(095)l.}(X_4)$	0.0268	-0.0011	0.0079	0.0156	$0.0733 = C_{h(0.95)..}(X_4)$
$C_{h(0.95)lW}(Y)$	0.0036	0.0132	0.0114	0.0174	$0.0436 = C_{h(0.95).W}(Y)$
$C_{h(0.95)lB}(Y)$	0.2243	0.0116	-0.0312	-0.0431	$0.1616 = C_{h(0.95).B}(Y)$
$C_{h(0.95)l.}(Y)$	0.2279	0.0248	-0.0198	-0.0257	$0.2072 = G_{h(0.95)}(Y)$

TABLE 29. - *Joint decomposition of* $G(Y) = 0.35$. *Contributions of within and between components for four subpopulations and four income sources. Calculation based on Bank of Italy 2014 sample survey data*

	$D=1$	$D=2$	$D=3$	$D \geq 4$	Italy
	$l=1$	$l=2$	$l=3$	$l=4$	
$C_{.lW}(X_1)$	0.0037	0.0051	0.0054	0.0116	$0.0258 = C_{..W}(X_1)$
$C_{.lB}(X_1)$	0.1002	0.0491	-0.0024	-0.0212	$0.1257 = C_{..B}(X_1)$
$C_{.l.}(X_1)$	0.1039	0.0542	0.0029	-0.0096	$0.1514 = C_{...}(X_1)$
$C_{.lW}(X_2)$	0.0031	0.0074	0.0024	0.0016	$0.0145. = C_{..W}(X_2)$
$C_{.lB}(X_2)$	0.0240	-0.0267	0.0128	0.0328	$0.0429 = C_{..B}(X_2)$
$C_{.l.}(X_2)$	0.0270	-0.0193	0.0152	0.0343	$0.0572 = C_{...}(X_2)$
$C_{.lW}(X_3)$	0.0028	0.0034	0.0030	0.0044	$0.0136 = C_{..W}(X_3)$
$C_{.lB}(X_3)$	0.0327	0.0163	0.0035	0.0019	$0.0544 = C_{..B}(X_3)$
$C_{.l.}(X_3)$	0.0355	0.0197	0.0066	0.0063	$0.0681 C_{...}(X_3)$
$C_{.lW}(X_4)$	0.0048	0.0058	0.0026	0.0039	$0.0171 = C_{..W}(X_4)$
$C_{.lB}(X_4)$	0.0278	0.0081	0.0082	0.0120	$0.0561 = C_{..B}(X_4)$
$C_{.l.}(X_4)$	0.0326	0.0139	0.0108	0.0160	$0.0733 = C_{...}(X_4)$
$C_{.lW}(Y)$	0.0143	0.0218	0.0134	0.0215	$0.0710 = C_{..W}(Y)$
$C_{.lB}(Y)$	0.1847	0.0467	0.0221	0.0255	$0.2790 = C_{..B}(Y)$
$C_{.l.}(Y)$	0.1990	0.0685	0.0355	0.0470	$0.3500 = G(Y)$

7.2.1 *Subpopulations contributions to the point* $G_h(Y)$ *and the synthetic* $G(Y)$ *indexes*

Starting from the $k \times k$ contributions in (44), we can obtain other decompositions of $G_h(Y)$. In particular, let:

$$C_{hl.}(Y) = \sum_{g=1}^{k} C_{hlg}(Y)$$

$$= \sum_{g=1}^{k} \frac{M_g(Y) - \bar{M}_{hl}(Y)}{M(Y)} \cdot \frac{n_{.g}}{N} \cdot p(l|h) \cdot 2p_{h.} - \sum_{g=1}^{k} \frac{M_g(Y) - y_h}{M(Y)} \cdot \frac{n_{.g}}{N} \cdot f(l|h) \cdot \frac{n_{h.}}{N}.$$

Using in this latter expression the relation (37) gives:

$$C_{hl.}(Y) = \frac{M(Y) - \bar{M}_{hl}(Y)}{M(Y)} \cdot p(l|h) \cdot 2p_{h.} - \frac{M(Y) - y_h}{M(Y)} \cdot f(l|h) \cdot \frac{n_{h.}}{N}. \quad (73)$$

Now putting (73) in (44) gives the following k additive decomposition of $G_h(Y)$:

$$G_h(Y) = \sum_{l=1}^{k} C_{hl.}(Y); \quad (74)$$

$C_{hl.}(Y)$ can be interpreted as the contribution of the subpopulation l to the point inequality index $G_h(Y)$. For more details on this point see: Zenga (2016a); Zenga and Valli (2016); Zenga (2016b).

Putting (74) in $G(Y) = \sum_{h=1}^{r} G_h(Y) \cdot \frac{n_{h.}}{N}$ gives

$$G(Y) = \sum_{h=1}^{r}\sum_{l=1}^{k} C_{hl.}(Y) \cdot \frac{n_{h.}}{N} = \sum_{l=1}^{k} C_{.l.}(Y); \tag{75}$$

$$C_{.l.}(Y) = \sum_{h=1}^{r} C_{hl.}(Y) \cdot \frac{n_{h.}}{N} \tag{76}$$

is the contribution of subpopulation l to $G(Y)$. It is important to remark that, by the use of the relations (38) and (42) in (11) we can obtain, in a more direct way, the decomposition (74).

$$G_h(Y) = \frac{M(Y) - \sum_{l=1}^{k} \bar{M}_{hl}(Y) \cdot p(l|h)}{M(Y)} \cdot 2p_{h.} - \frac{M(Y) - \sum_{l=1}^{k} y_h \cdot f(l|h)}{M(Y)} \cdot \frac{n_{h.}}{N} =$$

$$= \sum_{l=1}^{k}\left[\frac{M(Y) - \bar{M}_{hl}(Y)}{M(Y)} \cdot p(l|h) \cdot 2p_{h.} - \frac{M(Y) - y_h}{M(Y)} \cdot f(l|h) \cdot \frac{n_{h.}}{N}\right] = \sum_{l=1}^{k} C_{hl.}(Y).$$

Tables 26, 27, and 28 report the four subpopulations contributions $C_{hl.}(Y)$ to the three Gini point measures. For the interpretation of $C_{hl.}\{Y\}$ it is useful the following relation:

$$C_{hl.}(Y) = V_{hl.}(Y) \cdot 2p_{h.} - E_{hl.}(Y), \tag{77}$$

where

$$V_{hl.}(Y) = \frac{M(Y) - \bar{M}_{hl}(Y)}{M(Y)} \cdot p(l|h) \tag{78}$$

is the contribution of subpopulation l to the Bonferroni point index $V_h(Y)$, and

$$E_{hl.}(Y) = E_h(Y) \cdot f(l|h). \tag{79}$$

Note that, in (79),

$$E_h(Y) = \frac{M(Y) - y_h}{M(Y)} \cdot \frac{n_{h.}}{N} = A_h(Y) \cdot \frac{n_{h.}}{N} \cdot \tag{80}$$

In Section 7.2 it is shown that the three values of $E_h(Y)$ are negligible. Consequently, the corresponding values of $E_{hl.}(Y)$ are also negligible because the relative frequencies $f(l|h) = \frac{n_{hl}}{n_{h.}}$ are such that $\{0 \leq f(l|h) \leq 1\}$.

In conclusion, in the present application, the values of the contributions $C_{hl.}(Y)$ are very close to the product $V_{hl.}(Y) \cdot 2p_{h.}$.

Now, we illustrate the decomposition of each of the three Gini point indexes into the four subpopulations contributions.

For $G_{h(0.1)}(Y) = 0.1568$ we have:
$E_h(Y) = A_h(Y) \cdot \frac{n_{h.}}{N} = 0.65929 \cdot 0.000969 = 0.00063908; f(1|h) = 0.5510$, $f(2|h) = 0.2852$, $f(3|h) = 0.1127$, $f(4|h) = 0.0511$. Table 30 reports the calculation of the four contributions $C_{h(0.1)l.}(Y)$.

TABLE 30. - *Calculation of* $C_{h(0.1)l.}(Y) = V_{hl.}(Y) \cdot 0.2 - E_{hl.}(Y)$

l	$\frac{M(Y)-\bar{M}_{hl}(Y)}{M(Y)}$	$p(l\|h)$	$V_{hl.}(Y)$	$V_{hl.}(Y) \cdot 0.2$	$E_{hl.}(Y)$	$C_{h(0.1)l.}(Y)$
1	0.7696	0.5556	0.42759	0.08552	0.000352	0.085166
2	0.7809	0.1538	0.12010	0.02402	0.000188	0.023832
3	0.8270	0.1217	0.100645	0.02013	0.000072	0.020058
4	0.8171	0.1689	0.138008	0.027602	0.000033	0.027569
Tot.		1.0000	0.7864= $V_h(Y)$	0.15728= $V_h(Y) \cdot 0.2$	0.000639= $E_h(Y)$	0.1568= $G_h(Y)$

Table 30 shows that the relative variations of the lower means $\bar{M}_{hl}(Y)$ of the four subpopulations w.r.t. the mean $M(Y)$ are similar and very close to the value 0.7864) of the Bonferroni point index $V_{h(0.1)}(Y)$, while their relative weights $p(l|h)$ are very different. This explains why there are so remarkable differences among the Bonferroni contributions $V_{hl.}(Y)$ and among the Gini contributions $C_{h(0.1)l.}(Y)$. In particular, we note that the ‘‘number’’ of households with one components and with $Y \leq y_{h(0.1)} = 10400$ Euro is the 55.56% of ‘‘the number’’ of the households of the corresponding whole lower group. This explain why the greatest contribution to the Gini point index comes from the households with one component.

For $G_{h(0.5)}(Y) = 0.4813$ we have:
$E_h(Y) = A_h(Y) \cdot \frac{n_{h.}}{N} = 0.17747 \cdot 0.0001463 = 0.000026$; $f(1|h)$=1.00, $f(2|h) = f(3|h) = f(4|h) = 0.00$. Table 31 reports the calculation of the four contributions $C_{h(0.5)l.}(Y)$.

TABLE 31. - *Calculation of* $C_{h(0.5)l.}(Y) = V_{hl.}(Y) \cdot 1.0 - E_{hl.}(Y)$

l	$\frac{M(Y)-\bar{M}_{hl}}{M(Y)}$	$p(l\|h)$	$V_{hl.}(Y)$	$V_{hl.}(Y) \cdot 1.0$	$E_{hl.}(Y)$	$C_{h(0.5)l.}(Y)$
1	0.5126	0.4723	0.2421	0.2421	0.000026	0.242074
2	0.4400	0.2336	0.10278	0.10278	0.00	0.10278
3	0.4684	0.1252	0.05864	0.05864	0.00	0.05864
4	0.4600	0.1689	0.07769	0.07769	0.00	0.07769
Tot.		1.00	0.481214= $V_h(Y)$	0.481214= $V_h(Y) \cdot 1.0$	0.000026= $E_h(Y)$	0.4812= $G_h(Y)$

Table 31 shows that: the differences, between the four relative variations $(M(Y) - \bar{M}_{hl}(Y))/M(Y)$ and the value (0.4812) of the Bonferroni point index $V_{h(0.5)}(Y)$, are not remarkable; the relative weight $p(1|h)$ of the households with one component is still greater than those of the other subpopulations. Thus, the greatest contributions to the Bonferroni point index and to the Gini point index comes from the households with one component.

Finally, for $G_{h(0.95)}(Y) = 0.2072$ we have:
$E_h(Y) = A_h(Y) \cdot \frac{n_{h.}}{N} = -1.209193 \cdot 0.000137 = -0.0001656$; $f(2|h)$=1.00, $f(1|h) = f(3|h) = f(4|h) = 0.00$. Table 32 reports the calculation of the four contributions $C_{h(0.95)l.}(Y)$.

TABLE 32. - *Calculations of* $C_{h(0.95)l.}(Y) = V_{hl.}(Y) \cdot 1.9 - E_{hl.}(Y)$

l	$\frac{M(Y)-\bar{M}_{hl}}{M(Y)}$	$p(l\|h)$	$V_{hl.}(Y)$	$V_{hl.}(Y) \cdot 1.9$	$E_{hl.}(Y)$	$C_{h(0.95)l.}(Y)$
1	0.391951	0.3060	0.11994	0.22788	0.00	0.22788
2	0.047166	0.2752	0.01298	0.024662	-0.000165	0.024827
3	-0.05534	0.1881	-0.01041	-0.019777	0.00	-0.01977
4	-0.05873	0.2307	-0.013549	-0.02574	0.00	-0.02574
Tot.		1.0000	0.108962= $V_h(Y)$	0.207025= $V_h(Y) \cdot 1.9$	-0.000165= $E_h(Y)$	0.2072= $G_h(Y)$

Table 32 shows that there are remarkable differences among the four relative variations $(M(Y) - \bar{M}_{hl}(Y))/M(Y)$. In particular: the relative variation $(M(Y) - \bar{M}_{h1}(Y))/M(Y) = 0.3919506$ is a lot greater than those of the other three subpopulations; the relative variations $(M(Y) - \bar{M}_{h3}(Y))/M(Y)$ and $(M(Y) - \bar{M}_{h4}(Y))/M(Y)$ are negative. Moreover, the relative wheight of the households with one component: is greater than the one of the households with two components, and is a lot greater than those of the other two subpopulations. As a result, for $p = 0.95$, we have:

$$0.22788 = C_{h(0.95)1.}(Y) > G_{h(0.95)}(Y) = 0.2072.$$

In the last row of Table 29 are reported the contributions $C_{.l.}(Y)$ of each subpopulation to the synthetic Gini index $G(Y) = 0.35$.

TABLE 33. - *Contributions of each subpopulation to the three Gini point indexes and to the Gini synthetic index*

l	$C_{h(0.1)l.}(Y)$	$C_{h(0.5)l.}(Y)$	$C_{h(0.95)l.}(Y)$	$C_{.l.}(Y)$
1	0.08517	0.24207	0.22788	0.19900
2	0.02383	0.10278	0.02483	0.0685
3	0.02006	0.05864	-0.01977	0.0355
4	0.02757	0.07769	-0.02574	0.0470
Tot.	0.1568= $G_{h(0.1)}(Y)$	0.4812= $G_{h(0.5)}(Y)$	0.2072= $G_{h(0.95)}(Y)$	0.3500= $G(Y)$

Finally, Table 33 reports the contributions of each subpopulation to the three Gini point indexes and to the Gini synthetic index. The values reported in Table 33 and (76) help to understand why the greatest contribution to $G(Y)$ comes from the househols with one component.

7.2.2 *Within and between parts of the subpopulations contributions, and of the points and the synthetic indexes*

The contribution $C_{hl.}(Y) = \sum_{g=1}^{k} C_{hlg}(Y)$ can be split into a within $C_{hlW}(Y) = C_{hll}(Y)$ and a between $C_{hlB}(Y) = \sum_{g \neq l} C_{hlg}(Y)$ component:

$$C_{hl.}(Y) = C_{hlW}(Y) + C_{hlB}(Y). \tag{81}$$

In particular, using (45) in (81) gives:

$$C_{hlW}(Y) = \frac{M_l(Y) - \bar{M}_{hl}(Y)}{M(Y)} \cdot \frac{n_{.l}}{N} \cdot p(l/h) \cdot 2p_{h.} - \frac{M_l(Y) - y_h}{M(Y)} \cdot \frac{n_{.l}}{N} \cdot f(l/h) \cdot \frac{n_{h.}}{N}, \tag{82}$$

and

$$C_{hlB}(Y) = \sum_{g \neq l} \left(\frac{M_g(Y) - \bar{M}_{hl}(Y)}{M(Y)} \frac{n_{.g}}{N} \cdot p(l/h) 2p_{h.} - \frac{M_g(Y) - y_h}{M(Y)} \frac{n_{.g}}{N} f(l/h) \frac{n_{h.}}{N} \right). \tag{83}$$

Putting (81) in (74) gives:

$$G_h(Y) = \sum_{l=1}^{k} [C_{hlW}(Y) + C_{hlB}(Y)] = C_{h.W}(Y) + C_{h.B}(Y), \tag{84}$$

where

$$C_{h.W}(Y) = \sum_{l=1}^{k} C_{hlW}(Y) = \sum_{l=1}^{k} C_{hll}(Y), \tag{85}$$

and

$$C_{h.B}(Y) = \sum_{l=1}^{k} C_{hlB}(Y) = \sum_{l=1}^{k} \sum_{g \neq l} C_{hlg}(Y) \tag{86}$$

are respectively the within and the between parts of the Gini point index $G_h(Y)$.

Finally, putting (81) in (76) gives:

$$C_{.l.}(Y) = C_{.lW}(Y) + C_{.lB}(Y), \tag{87}$$

where

$$C_{.lW}(Y) = \sum_{h=1}^{r} C_{hlW}(Y) \cdot \frac{n_{h.}}{N} \text{ and } C_{.lB}(Y) = \sum_{h=1}^{r} C_{hlB}(Y) \cdot \frac{n_{h.}}{N} \tag{88}$$

are respectively the within and the between part of the contribution $C_{.l.}(Y)$ of subpopulation l to the synthetic Gini index $G(Y)$.

In conclusion, putting (87) in (75) gives:

$$G(Y) = C_{..W}(Y) + C_{..B}(Y), \tag{89}$$

where

$$C_{..W}(Y) = \sum_{l=1}^{k} C_{.lW}(Y) \text{ and } C_{..B}(Y) = \sum_{l=1}^{k} C_{.lB}(Y) \tag{90}$$

The last three rows of Tables 26, 27, and 28 report the within and the between parts of the subpopulations contributions, as well as the within and the between parts of the Gini point indexes. All the values of the reported within parts are positive. The values of the between parts for $p = 0.1$ and $p = 0.5$ are positive, while for $p = 0.95$ two values are positive and two values are negative. Table 34 illustrates the calculation to obtain for $p = 0.95$ the four values $V_{h4g}(Y) \cdot 2 \cdot 0.95$, $(g = 1, 2, 3, 4)$; these values are very close to the corresponding Gini's contributions $C_{h4g}(Y)$. Moreover, Table 34 gives: the sum $\sum_{g=1}^{3} V_{h4g}(Y) \cdot 1.9\ C_{h4B}(Y)$ and $V_{h44}(Y) \cdot 1.9 \simeq C_{h44}(Y) = C_{h4W}(Y)$.

TABLE 34. - *Calculation of the within and the between parts of the contribution* $C_{h(0.95)4.}(Y)$

g	(a) $\frac{M_g(Y)-\bar{M}_{h4}(Y)}{M}$	(b) $\frac{n_{.g}}{N}$	$(c) =$ $(a) \cdot (b)$	$(d) =$ $p(4\|h)$	$V_{h4g} =$ $(c) \cdot (d)$	$C_{h4g}(Y) \simeq$ $V_{h4g}(Y) \cdot 1.9$
1	$\frac{19207.1-32318.0}{30525.2} =$ -0.42951	0.2933	-0.125975	0.2307	-0.02906	-0.05522
2	$\frac{31895.2-\bar{M}_{h4}}{M} =$ -0.01385	0.2730	-0.003781	0.2307	-0.000872	-0.001657
3	$\frac{37265.4-\bar{M}_{h4}}{M} =$ 0.162077	0.1935	0.031362	0.2307	0.007235	0.013746
					$V_{h4B}(Y) =$ -0.022697	$V_{h4B}(Y) \cdot 1.9 =$ $-0.0431 \simeq C_{h4B}$
4	$\frac{37358.8-\bar{M}_{h4}(Y)}{M(Y)} =$ 0.165136	0.2402	0.039666	0.2307	0.00915= $V_{h44=}V_{h4W}$	$V_{h4W} \cdot 1.9 =$ $0.01739 \simeq C_{h4W}$

Note that, the relative differences $\{(M_1(Y) - \bar{M}_{h4}(Y))/M(Y)\} = -0.4295$ and $\{(M_2(Y) - \bar{M}_{h4}(Y))/M(Y)\} = -0.0183$ are negative because $32318 = \bar{M}_{h4} > M_1 = 19207.1$ and $32318 = \bar{M}_{h4} > M_2 = 31895.2$.

The other two relative differences $\{(M_3(Y) - \bar{M}_{h4}(Y))/M(Y)\} = 0.1621$ and $\{(M_4(Y) - \bar{M}_{h4}(Y))/M(Y)\} = 0.1651$ are positive because M_3 (Y) and $M_4(Y)$ are greater than $\bar{M}_{h4}(Y)$. We have to point out that the within component can't be negative because $(M_l(Y) \geq \bar{M}_{hl}(Y)), \forall(l, h)$.

The last three rows of Table 29 report the within and the between parts of the subpopulation contributions, as well as the within and the between parts of the synthetic Gini index $G(Y) = 0.3500$.

Table 35 reports the shares of the within part in the (three) points and in the synthetic Gini indexes, and in the (corresponding) subpopulation contributions.

TABLE 35. - *Within part shares: in the three point indexes* $[C_{h(p).W}(Y)/G_{h(p)}(Y)]$, *in the synthetic Gini index* $[C_{..W}(Y)/G(Y)]$, *and inthe subpopulation contributions* $[C_{h(p)lW}(Y)/C_{h(p)l.}(Y)]$ *and* $[C_{.lW}(Y)/C_{.l.}(Y)]$

	$l=1$	$l=2$	$l=3$	$l=4$	Italy
$\frac{C_{h(0.1)lW}(Y)}{C_{h(0.1)l.}(Y)}$	0.1525	0.2887	0.2439	0.3043	$0.2117 = \frac{C_{h(0.1).W}(Y)}{G_{h(0.1)}(Y)}$
$\frac{C_{h(0.5)lW}(Y)}{C_{h(051)l.}(Y)}$	0.0814	0.3000	0.2845	0.3578	$0.1976 = \frac{C_{h(0.5).W}(Y)}{G_{h(0.5)}(Y)}$
$\frac{C_{h(0.95)lW}(Y)}{C_{h(0.95)l.}(Y)}$	0.0158	0.5322	-0.5758	-0.6770	$0.2040 = \frac{C_{h(0.95).W}(Y)}{G_{h(0.95)}(Y)}$
$\frac{C_{.lW}(Y)}{C_{.l.}(Y)}$	0.0719	0.3189	0.3775	0.4574	$0.2029 = \frac{C_{..W}(Y)}{G(Y)}$

In the present application the within part: is the 20.29% of the synthetic Gini index, and takes similar percentages in the three point indexes.Vice versa, in the subpopulations contributions the shares take on very different percentages. The shares of the within part in the subpopulation contributions are, in the present application, very close to:

$$\frac{C_{hlW}(Y)}{C_{hl.}(Y)} \simeq \frac{M_l(Y) - \bar{M}_{hl}(Y)}{M(Y) - \bar{M}_{hl}(Y)} \cdot \frac{n_{.l}}{N}. \tag{91}$$

For $M_l(Y) > M(Y)$, the difference $[M(Y) - \bar{M}_{hl}(Y)]$ can assume negative values. This explains why, for $p = 0.95$ and $l = 3$ and $l = 4$, the corresponding ratios are negative.

7.2.3 *Joint decompositions by four subpopulations and four sources of* $G_h(Y)$ *and* $G(Y)$

Tables 26, 27, and 28 report the decompositions by the four sources (X_1, X_2, X_3, X_4) of: the contributions $C_{h(p)l.}(Y) = \sum_{j=1}^{4} C_{h(p)l.}(X_j)$ and of the point indexes $G_{h(p)}(Y) = \sum_{j=1}^{4} C_{h(p)..}(X_j)$. Table 29 reports the decompositions by the four sources of: the contributions $C_{.l.}(Y) = \sum_{j=1}^{4} C_{.l.}(X_j)$ and of the synthetic Gini index $G(Y)=\sum_{j=1}^{4} C_{...}(X_j)$.

The relative contribution of X_j to:

1. the contribution $C_{h(p)l.}(Y)$ is given by

$$\lambda_{h(p)l.}(X_j) = \left[C_{h(p)l.}(X_j)/C_{h(p)l.}(Y)\right] \qquad (l = 1, ..., 4) \tag{92}$$

2. the point index $G_{h(p)}(Y)$ is given by

$$\lambda_{h(p)..}(X_j) = \left[C_{h(p)..}(X_j)/G_{h(p)}(Y)\right]; \tag{93}$$

3. the contributions $C_{.l.}(Y)$ is given by

$$\lambda_{.l.}(X_j) = \left[C_{.l.}(X_j)/C_{.l.}(Y)\right], \qquad (l = 1, ..., 4); \tag{94}$$

4. the synthetic index $G(Y)$ is given by

$$\lambda_{...}(X_j) = \left[C_{...}(X_j)/G(Y)\right]. \tag{95}$$

Note that: between the shares $\lambda_{h(p)..}(X_j)$ and the corresponding k shares $\lambda_{h(p)l.}(X_j)$ there is the following relation

$$\lambda_{h(p)..}(X_j) = \sum_{l=1}^{k} \lambda_{h(p)l.}(X_j) \cdot \frac{C_{h(p)l.}(Y)}{G_{h(p)}(Y)}, \qquad (j = 1, ..., 4); \tag{96}$$

while between the shares $\lambda_{...}(X_j)$ and the corresponding k shares $\lambda_{.l.}(X_j)$ there is the following relation

$$\lambda_{...}(X_j) = \sum_{l=1}^{k} \lambda_{.l.}(X_j) \cdot \frac{C_{.l..}(Y)}{G(Y)}, \qquad (j = 1, ..., 4). \tag{97}$$

In other words, the share $\lambda_{h(p)..}(X_j)$ is the weighted arithmetic mean of the corresponding shares $\lambda_{h(p)l.}(X_j)$ with wheights $\left[C_{h(p)l.}(Y)/G_{h(p)}(Y)\right]$, and the share $\lambda_{...}(X_j)$ is the weighted arithmetic mean of the corresponding shares $\lambda_{.l.}(X_j)$ with wheights $[C_{.l..}(Y)/G(Y)]$.

Table 36 reports the relative contributions: $\lambda_{h(p)l.}(X_j)$, $\lambda_{h(p)..}(X_j)$, $\lambda_{.l.}(X_j)$, $\lambda_{...}(X_j)$. In the case of the Joint decomposition of $G_{h(p)}(Y)$ it is worth to check whether the shares $\lambda_{h(p)l.}(X_j)$ are influenced by the subpopulations, and by the corresponding cumulative frequency p; while in the case of the decomposition of $G(Y)$ it is worth to check if the shares are influenced by the subpopulations.

For $p = 0.10$ the four subpopulations have a limited influence on the values of the shares $\lambda_{h(0.10)l.}(X_j)$. For $p = 0.5$ the subpopulations have a great influence on the shares $\lambda_{h(0.50)l.}(X_j)$. For $p = 0.95$ there are many negative shares. Note that in the case of the sinthetic index there are two negative shares, too. We note also that the shares $\lambda_{h(p)l.}(X_j)$ and $\lambda_{h(p)..}(X_j)$, and the shares $\lambda_{.l.}(X_j)$ and $\lambda_{...}(X_j)$ reported in Table 36 are coherent with relation (96) and (97), respectively.

TABLE 36. - *Relative contributions:* $\lambda_{h(p)l.}(X_j) = \left[C_{h(p)l.}(X_j)/C_{h(p)l.}(Y)\right]$ $\lambda_{h(p)..}(X_j) = \left[C_{h(p)..}(X_j)/G_{h(p)}(Y)\right], \lambda_{.l.}(X_j) = \left[C_{.l.}(X_j)/C_{.l.}(Y)\right], \lambda_{...}(X_j) = \left[C_{...}(X_j)/G(Y)\right]$

.	$l = 1$	$l = 2$	$l = 3$	$l = 4$	Italy
$p = 0.10; h = 440$					
$\lambda_{hl.}(X_1)$	0.4327	0.4720	0.4314	0.4037	$0.4344 = \lambda_{h..}(X_1)$
$\lambda_{hl.}(X_2)$	0.2289	0.2092	0.2532	0.2695	$0.2361 = \lambda_{h..}(X_2)$
$\lambda_{hl.}(X_3)$	0.1387	0.1147	0.1318	0.1396	$0.1332 = \lambda_{h..}(X_3)$
$\lambda_{hl.}(X_4)$	0.1982	0.2031	0.1852	0.1871	$0.1956 = \lambda_{h..}(X_4)$
	1	1	1	1	1
$p = 0.50; h = 3262$					
$\lambda_{hl.}(X_1)$	0.5073	0.5936	0.3047	0.1774	$0.4478 = \lambda_{h..}(X_1)$
$\lambda_{hl.}(X_2)$	0.1412	-0.0213	0.3010	0.4253	$0.1718 = \lambda_{h..}(X_2)$
$\lambda_{hl.}(X_3)$	0.1751	0.2032	0.1551	0.1416	$0.1733 = \lambda_{h..}(X_3)$
$\lambda_{hl.}(X_4)$	0.1765	0.2245	0.2392	0.2557	$0.2071 = \lambda_{h..}(X_4)$
	1	1	1	1	1
$p = 0.95; h = 6945$					
$\lambda_{hl.}(X_1)$	0.5805	3.1660	2.1890	4.2730	$0.2780 = \lambda_{h..}(X_1)$
$\lambda_{hl.}(X_2)$	0.1062	-3.3846	-0.7481	-2.7946	$0.1299 = \lambda_{h..}(X_2)$
$\lambda_{hl.}(X_3)$	0.1959	1.2619	-0.042	0.1288	$0.3546 = \lambda_{h..}(X_3)$
$\lambda_{hl.}(X_4)$	0.1175	-0.044	-0.3988	-0.6072	$0.2374 = \lambda_{h..}(X_4)$
	1	1	1	1	1
Synthetic Gini index					
$\lambda_{.l.}(X_1)$	0.5219	0.7912	0.082	-0.2034	$0.4325 = \lambda_{...}(X_1)$
$\lambda_{.l.}(X_2)$	0.1358	-0.281	0.4281	0.7301	$0.1637 = \lambda_{...}(X_2)$
$\lambda_{.l.}(X_3)$	0.1785	0.2876	0.1855	0.1338	$0.1945 = \lambda_{...}(X_3)$
$\lambda_{.l.}(X_4)$	0.1640	0.2024	0.3045	0.3395	$0.2093 = \lambda_{...}(X_4)$
	1	1	1	1	1

7.3 *Decomposition by four sources, for each of the four subpopulations and for the whole country, of the (corresponding) point and synthetic Gini indexes*

Table 37 reports, for each of the four subpopulations l and for the whole country, the decompositions by the four sources (X_1, X_2, X_3, X_4), of the corresponding (three) Gini point and synthetic indexes.

- For $p = 0,10$: the greatest index $G_{h(0.10)l}(Y) = 0.239$ comes from the households with $D = 1$ component, the lowest $G_{h(0.10)l}(Y) = 0.088$ comes from the households with $D = 2$ components, and for the whole Country $G_{h(0,10)}(Y) = 0.1568$.
- For $p = 0,50$: the greatest index $G_{h(0.50)l}(Y) = 0.397$ comes from the households with $D = 2$ components, the lowest $G_{h(0.50)l}(Y) = 0.363$ comes from the households with $D = 1$ component, and for the whole Country $G_{h(0,50)}(Y) = 0.4813$.
- For $p = 0,95$: the greatest index $G_{h(0.95)l}(Y) = 0.25$ comes from the households with $D = 3$ components, the lowest $G_{h(0.95)l}(Y) = 0.067$ comes from the households with $D = 1$ component, and for the whole Country $G_{h(0,95)}(Y) = 0.2072$.

Finally, the greatest synthetic index $G_{.l}(Y) = 0.327$ comes from the households with $D \geq 4$ components, the lowest one $G_{.l}(Y) = 0.3055$ comes comes from the households with $D = 1$ component, and for the whole Country $G(Y) = 0.35$. Table 38 reports the relative contributions of X_j:

- $\lambda_{h(p)l}(X_j) = \frac{C_{h(p)l}(X_j)}{G_{h(p)l}(Y)}$, $l = 1, ..., 4$; $p = 0.10, 0.50, 0.95$, to the point index $G_{h(p)l}(Y)$ of the subpopulation l;
- $\lambda_{h(p)..}(X_j) = \frac{C_{h(p)..}(X_j)}{G_{h(p)}(Y)}$, $p = 0.10, 0.50, 0.95$, to the point index $G_{h(p)}(Y)$ of the whole Country;
- $\lambda_{.l}(X_j) = \frac{C_{.l}(X_j)}{G_{.l}(Y)}$, $l = 1, ..., 4$, to the synthetic index $G_{.l}(Y)$ of the subpopulation l;
- $\lambda_{...}(X_j) = \frac{C_{...}(X_j)}{G(Y)}$, to the synthetic index $G_{.}(Y)$ of the whole population.

Table 38 shows that the relative contributions (to the inequality) of the variates X_1 (payroll income) and X_3 (net self employment income) are greater for the households with $(D = 3)$ and $(D \geq 4)$ w.r.t. the ones with $(D = 1)$ and $(D = 2)$. Vice versa, for the variate X_2 (pensions and net transfers) and X_4 (property incomes) there is a reverse relation. Table 38 shows also that there are important differences between the relative contributions of the subpopulations and the ones of the whole Country.

It is worth to observe that there are also very remarkable differences between the relative contributions $\lambda_{h(p)l.}(X_j) = [C_{h(p)l.}(X_j)/C_{h(p)l.}(Y)]$ of Table 36 and the relative contributions $\lambda_{h(p)l}(X_j) = [C_{h(p)l}(X_j)/G_{h(p)l}(Y)]$ of Table 38. To understand why there are these ''remarkable'' differences we observe, first of all, that the contribution $\lambda_{hl.}(X_j)$ is also equal to:

$$\begin{aligned} \lambda_{hl.}(X_j) &= \frac{C_{hl.}(X_j)}{C_{hl.}(Y)} = \frac{C_{hlW}(X_j) + C_{hlB}(X_j)}{C_{hlW}(Y) + C_{hlB}(Y)} \\ &= \frac{C_{hlW}(X_j)}{C_{hlW}(Y)} \cdot \frac{C_{hlW}(Y)}{C_{hl.}(Y)} + \frac{C_{hlB}(X_j)}{C_{hlB}(Y)} \cdot \frac{C_{hlB}(Y)}{C_{hl.}(Y)}. \end{aligned} \tag{98}$$

Now in the present application: from formula (67) it derives that the point inequality index $G_{hl}(Y)$ of subpopulation l is very close to $[(M_l(Y) - \bar{M}_{hl}(Y))/M_l(Y)] \cdot 2 \cdot p_{hl}$, and from formula (73) it derives that the contribution $C_{hl}(X_j)$ is very close to

$\left[(M_l(X_j) - \bar{M}_{hl}(X_j))/M_l(Y)\right] \cdot 2 \cdot p_{hl}$. Thus, the value of relative contributions $\lambda_{hl}(X_j)$of X_j to the point index $G_{hl}(Y)$ of the subpopulation l, can be approximated by the ratio

$$\lambda_{hl}(X_j) \approx \frac{\left[(M_l(X_j) - \bar{M}_{hl}(X_j))/M_l(Y)\right]2p_{hl}}{[(M_l(Y) - \bar{M}_{hl}(Y))/M_l(Y)]2p_{hl}} = \frac{M_l(X_j) - \bar{M}_{hl}(X_j)}{M_l(Y) - \bar{M}_{hl}(Y)}.$$

Moreover, in the present application, the value of $C_{hlW}(Y) = C_{hll}(Y)$ is very close to $[(M_l(Y) - \bar{M}_{hl}(Y))/M(Y)] \cdot \frac{n_{.l}}{N} \cdot p(l/h) \cdot 2p_{h.}$ (see formula 82), and the value of $C_{hlW}(X_j) = C_{hll}(X_j)$ is very close to $\left[(M_l(X_j) - \bar{M}_{hl}(X_j))/M(Y)\right] \cdot \frac{n_{.l}}{N} \cdot p(l/h) \cdot 2p_{h.}$ (see formula 64). Thus, the value of the ratio

$$\frac{C_{hlW}(X_j)}{C_{hlW}(Y)} = \frac{C_{hll}(X_j)}{C_{hll}(Y)}$$

can also be approximated by the ratio

$$\frac{\left[(M_l(X_j) - \bar{M}_{hl}(X_j))/M(Y)\right] \cdot \frac{n_{.l}}{N} \cdot p(l|h) \cdot 2p_{h.}}{[(M_l(Y) - \bar{M}_{hl}(Y))/M(Y)] \cdot \frac{n_{.l}}{N} \cdot p(l|h) \cdot 2p_{h.}} = \frac{M_l(X_j) - \bar{M}_{hl}(X_j)}{M_l(Y) - \bar{M}_{hl}(Y)}.$$

Thus, in the present application $\lambda_{hl}(X_j) \approx \left[C_{hlW}(X_j)\right]/[C_{hlW}(Y)]$. Finally, using this approximation in (98) gives

$$\lambda_{hl.}(X_j) \approx \lambda_{hl}(X_j) \cdot \frac{C_{hlW}(Y)}{C_{hl.}(Y)} + \frac{C_{hlB}(X_j)}{C_{hlB}(Y)} \cdot \frac{C_{hlB}(Y)}{C_{hl.}(Y)}. \tag{99}$$

In conclusion, according to the relation (99) the contributions $\lambda_{hl.}(X_j)$ reported in Table 36 are the mean values of the corresponding $\lambda_{hl}(X_j)$ relative cointributions reported in Table 38 and the corresponding between ratios $\left[C_{hlB}(X_j)\right]/C_{hlB}(Y)$. Note that the contribution $C_{hlW}(X_j)$, $C_{hlW}(Y)$, $C_{hlB}(X_j)$, and $C_{hlB}(Y)$ are reported in Tables 26, 27, 28.

TABLE 37. - *Decomposition by sources of the point indexes $G_{hl}(Y)$ and $G_h(Y)$, and of the synthetic indexes $G_{.l}(Y)$ and $G(Y)$*

	$l=1$	$l=2$	$l=3$	$l=4$	Italy
	$C_{h1}(\cdot)$	$C_{h2}(\cdot)$	$C_{h3}(\cdot)$	$C_{h4}(\cdot)$	$C_{h..}(\cdot)$
$p=0.10; h=440$					
X_1	0.065	0.024	0.053	0.073	0.068
X_2	0.081	0.037	0.020	0.009	0.037
X_3	0.025	0.006	0.017	0.020	0.021
X_4	0.068	0.020	0.018	0.018	0.031
	0.239	0.088	0.107	0.119	0.157
	$G_{h1}(Y)$	$G_{h2}(Y)$	$G_{h3}(Y)$	$G_{h4}(Y)$	$G_h(Y)$
$p=0.5; h=3262$					
X_1	0.088	0.099	0.162	0.222	0.216
X_2	0.081	0.144	0.066	0.028	0.083
X_3	0.065	0.050	0.070	0.074	0.083
X_4	0.129	0.105	0.068	0.069	0.100
	0.363	0.397	0.365	0.393	0.481
	$G_{h1}(Y)$	$G_{h2}(Y)$	$G_{h3}(Y)$	$G_{h4}(Y)$	$G_h(Y)$
$p=0.95; h=6945$					
X_1	0.006	0.018	0.062	0.080	0.058
X_2	0.002	0.044	0.037	0.026	0.027
X_3	0.039	0.057	0.094	0.083	0.073
X_4	0.019	0.049	0.058	0.059	0.049
	0.067	0.169	0.250	0.246	0.207
	$G_{h1}(Y)$	$G_{h2}(Y)$	$G_{h3}(Y)$	$G_{h4}(Y)$	$G_h(Y)$
Decomposition by sources of the synthetic Gini indexes $G_{.l}(Y)$ and $G(Y)$					
	$C_{.1}(\cdot)$	$C_{.2}(\cdot)$	$C_{.3}(\cdot)$	$C_{.4}(\cdot)$	$C_{...}(\cdot)$
X_1	0.087	0.073	0.124	0.170	0.151
X_2	0.073	0.104	0.059	0.024	0.057
X_3	0.046	0.048	0.079	0.071	0.068
X_4	0.099	0.082	0.063	0.062	0.073
	0.306	0.308	0.325	0.327	0.350
	$G_{.1}(Y)$	$G_{.2}(Y)$	$G_{.3}(Y)$	$G_{.4}(Y)$	$G(Y)$

TABLE 38. - *Relative contributions* $\lambda_{hl}(X_j)$ *and* $\lambda_{h..}(X_j)$, *and* $\lambda_{.l}(X_j)$ *and* $\lambda_{...}(X_j)$

	$l = 1$	$l = 2$	$l = 3$	$l = 4$	Italy
	$\lambda_{h1}(\cdot)$	$\lambda_{h2}(\cdot)$	$\lambda_{h3}(\cdot)$	$\lambda_{h4}(\cdot)$	$\lambda_{h..}(\cdot)$
$p = 0.10$					
X_1	0.270	0.277	0.495	0.614	0.434
X_2	0.338	0.422	0.185	0.071	0.236
X_3	0.106	0.074	0.161	0.166	0.133
X_4	0.284	0.227	0.161	0.149	0.196
	1.00	1.00	1.00	1.00	1.00
$p = 0.5$					
X_1	0.243	0.249	0.442	0.565	0.448
X_2	0.222	0.362	0.180	0.071	0.172
X_3	0.179	0.125	0.192	0.188	0.173
X_4	0.356	0.264	0.186	0.176	0.207
	1.00	1.00	1.00	1.00	1.00
$p = 0.95$					
X_1	0.089	0.107	0.246	0.325	0.278
X_2	0.034	0.262	0.148	0.107	0.130
X_3	0.587	0.337	0.375	0.337	0.355
X_4	0.289	0.294	0.231	0.238	0.237
	1.00	1.00	1.00	1.00	1.00
Relative contributions $\lambda_{.l}(X_j)$ and $\lambda_{...}(X_j)$					
	$\lambda_{.1}(\cdot)$	$\lambda_{.2}(\cdot)$	$\lambda_{.3}(\cdot)$	$\lambda_{.4}(\cdot)$	$\lambda_{...}(\cdot)$
X_1	0.284	0.237	0.382	0.520	0.433
X_2	0.239	0.339	0.182	0.079	0.164
X_3	0.150	0.157	0.241	0.216	0.195
X_4	0.325	0.267	0.195	0.189	0.209
	1.00	1.00	1.00	1.00	1.00

8. CONCLUSIONS

Let, $X_1, \ldots, X_j, \ldots, X_c$ be non-negative variates (income sources) observable on each of the N units of the population, and $Y = \sum_{j=1}^{c} X_j$ be the total income. Let: $\{0 \le y_1 < \ldots < y_h < \ldots < y_r\}$ be the set of distinct values assumed by the variate Y and $\{n_{1.}, \ldots, n_{h.}, \ldots, n_{r.}; \sum_{h=1}^{r} n_{h.} = N\}$ be the corresponding frequencies.

In this paper the value of the synthetic Gini index $G(Y)$ is obtained by the following expression related to the Lorenz curve:

$$G(Y) = 2CA = 2\sum_{h=1}^{r} \frac{1}{2}[(p_{h.} - q_{h.}) + (p_{h-1.} - q_{h-1.})]\frac{n_{h.}}{N} = \sum_{h=1}^{r} G_h(Y)\frac{n_{h.}}{N},$$

where CA is the ''concentration area'' and

$$G_h(Y) = [(p_{h.} - q_{h.}) + (p_{h-1.} - q_{h-1.})]$$

is the point Gini index. In order to obtain the decompositions proposed in the present paper the latter expression of $G_h(Y)$ has been transformed in the following equivalent expression

$$G_h(Y) = \left\{\frac{M(Y) - \bar{M}_{h.}(Y)}{M(Y)} \cdot p_{h.} \cdot 2 - \frac{n_{h.}}{N} \cdot \frac{M(Y) - y_h}{M(Y)}\right\},$$

where $M(Y)$ is the mean of Y evaluated on the N units of the population and $\bar{M}_{h.}(Y)$ is the mean of Y evaluated on the $P_{h.} = \sum_{t=1}^{h} n_{t.}$ units of the lower group $\{Y \leq y_h\}$.

It is worth to point out that the decomposition proposed in this paper are obtained by the (recent) ''two step approach''. In particular, in the first step are obtained the decompositions by sources, by subpopulations, and the joint decomposition by subpopulations and sources of $G_h(Y)$, and in the second step the previous decompositions are extended to the synthetic Gini index $G(Y)$.

A very important result of the decomposition by subpopulations is that the value of the point index $G_h(Y)$ is decomposed in the sum of the contributions $C_{hl.}(Y)$ of each subpopulation $l : l = 1, \ldots, k$. Moreover, $C_{hl.}(Y)$ is decomposed in a within and a between part $C_{hl.}(Y) = C_{hlW}(Y) + C_{hlB}(Y)$; in the within parts only incomes of the same subpopulation l are compared, while in the between parts are compared only incomes of different subpopulations.

From the contribution $C_{hl.}(Y)$ we have obtained the following decomposition by sources: $C_{hl.}(Y) = \sum_{j=1}^{c} C_{hl.}(X_j)$, where $C_{hl.}(X_j)$ is the contribution of the variate X_j to $C_{hl.}(Y)$. For $C_{hl.}(X_j)$ we have also obtained the decomposition in a within and a between part: $C_{hl.}(X_j) = C_{hlW}(X_j) + C_{hlB}(X_j)$.

We remark that, in the second step we have obtained many additive decompositions of the synthetic Gini index $G(Y)$. In particular, $G(Y) = \sum_{l=1}^{k} C_{.l.}(Y)$, where $C_{.l.}(Y) = \sum_{h=1}^{r} C_{hl.}(Y) \cdot \frac{n_{h.}}{N}$ is the contribution of the subpopulation l to the synthetic Gini index.

In the present paper $G_{hl}(Y)$ denotes the point Gini inequality index of the subpopulation l, and the corresponding synthetic Gini inequality index (of the subpopulation l) is given by: $G_{.l}(Y) = \sum_{h=1}^{r} G_{hl}(Y) \cdot \frac{n_{hl}}{n_{.l}}$, where $n_{.l}$ is the size of the subpopulation l, and n_{hl} is the frequency of y_h in the subpopulation l. In the usual way we have also obtained for $G_{hl}(Y)$ the following decomposition by sources: $G_{hl}(Y) = \sum_{j=1}^{c} C_{hl}(X_j)$, where $C_{hl}(X_j)$ is the contribution of X_j to the point index $G_{hl}(Y)$.

The data used in the application of the present paper were supplied by the 2014 Central Bank of Italy sample survey on household income and wealth. In particular we deal with the household net disposable income Y, that is the sum of: payroll income X_1, pensions and net transfers X_2, net self employment income X_3, and property incomes X_4. Moreover, the 8156 households are partitioned into $k = 4$ subpopulations according to the number of persons D:$l = 1$ for households with $D = 1$; $l = 2$ for households with $D = 2$;$l = 3$ for households with $D = 3$; and $l = 4$ for households with $D \geq 4$. Table 39 reports for each subpopulation and for the whole country the relatve frequency $n_{.l}/N$, the mean $M_l(Y)$ and the synthetic Gini index $G_{.l}(Y)$. This Table shows that the values of the synthetic Gini index $G_{.l}(Y)$ of the households with $(D = 3)$ and $(D \geq 4)$ are a little bit greater than those of the househols with $(D = 1)$ and $(D = 2)$, and that the value of the synthetic Gini index $G(Y)$ of the whole country is greater than the corresponding index of the households with $(D \geq 4)$.

TABLE 39. - *Relatve frequency $n_{.l}/N$, mean $M_l(Y)$ and synthetic Gini index $G_{.l}(Y)$ for the application at hand*

	$D = 1$	$D = 2$	$D = 3$	$D \geq 4$	Italy
	$l = 1$	$l = 2$	$l = 3$	$l = 4$	
$n_{.l}/N$	0.2933	0.273	0.1935	0.2402	1
$M_l(Y)$	19207.1	31895.2	37265.4	37358.8	30525.2
$G_{.l}(Y)$	0.3055	0.3075	0.3254	0.3271	0.3500

In the paper are analyzed the decompositions of the synthetic index $G(Y) = 0.35$ and the decompositions of the point measure $G_{h(p)}(Y) = G_{(p)}(Y)$ for the following three values of the cumulative frequency p: $p = 0.10$, $p = 0.50$ and $p = 0.95$. In the present application, the values of the point Gini index are very close to the first term of the right side of the expression (11): $G_h(Y) \simeq \frac{M(Y)-\bar{M}_{h.}(Y)}{M(Y)} \cdot p_{h.} \cdot 2$. Consequently, for the interpretation of the contribution $C_{hl.}$ is useful the following approximation

$$C_{hl.}\{Y\} \simeq \frac{M(Y) - \bar{M}_{hl}(Y)}{M(Y)} \cdot p(l|h) \cdot 2p_{h.} = V_{hl.}(Y) \cdot 2p_{h.},$$

where $V_{hl.}(Y) = \frac{M(Y)-\bar{M}_{hl}(Y)}{M(Y)} \cdot p(l|h)$ is the contribution of subpopulation l to the Bonferroni point index, $\bar{M}_{hl}(Y)$ is the lower mean of the subpopulation l and $p(l|h) = P_{hl}/P_{h.}$ is the relative frequency of the subpopulation l in the lower group. Table 40 reports the contributions of each subpopulation to the three Gini point index and to the synthetic Gini index. In particular this Table show that: the greatest contribution to the points and to the synthetic indexes comes from the households with one component. It is important to remark that the contributions of the subpopulations $l = 3$ and $l = 4$ to the point index $G_{h(0.95)}(Y) = 0.2072$ are negative. The relation $C_{.l.}(Y) = \sum_{h=1}^{r} C_{hl.}(Y) \cdot \frac{n_{h.}}{N}$ explains why the greatest contribution to $G(Y)$ comes from the households with one component.

TABLE 40. - *Subpopulations contributions to the point and synthetic Gini index*

l	$C_{h(0.1)l.}(Y)$	$C_{h(0.5)l.}(Y)$	$C_{h(0.95)l.}(Y)$	$C_{.l.}(Y)$
1	0.08517	0.24207	0.22788	0.19900
2	0.02383	0.10278	0.02483	0.0685
3	0.02006	0.05864	-0.01977	0.0355
4	0.02757	0.07769	-0.02574	0.0470
Tot.	0.1568= $G_{h(0.1)}(Y)$	0.4812= $G_{h(0.5)}(Y)$	0.2072= $G_{h(0.95)}(Y)$	0.3500= $G(Y)$

For $p = 0.1$, Table 30 shows that the relative variations of the lower means $\bar{M}_{hl}(Y)$ of the four subpopulations w.r.t. the mean $M(Y)$ are similar and very close to the value 0.7864) of the Bonferroni point index $V_{h(0.1)}(Y)$, while their relative weights $p(l|h)$ are very different. This explains why there are so remarkable differences among the Gini contributions $C_{h(0.1)l.}(Y)$. In particular, we note that the ‘‘number’’ of households with one components and whit $Y \leq y_{h(0.1)} = 10400$ Euro is the 55.56% of ‘‘the number’’ of the households of the corresponding whole lower group. This explain why the greatest contribution to the Gini point index comes from the households with one component. For $p = 0.95$ Table 32 shows that the relative variation $(M(Y) - \bar{M}_{h3}(Y))/M(Y)$ and $M(Y) - \bar{M}_{h4}(Y))/M(Y)$ are negative and this explains why there are two negative contributions.

For the interpretation of the within and the beetween parts of the contribution $C_{hl.}(Y)$ the following approximations are useful

$$C_{hlW}(Y) \simeq \frac{M_l(Y) - \bar{M}_{hl}(Y)}{M(Y)} \cdot \frac{n_{.l}}{N} \cdot p(l|h) \cdot 2p_{h.}$$

$$C_{hlB}(Y) \simeq \sum_{g \neq l} \frac{M_g(Y) - \bar{M}_{hl}(Y)}{M(Y)} \frac{n_{.g}}{N} \cdot p(l|h) \cdot 2p_{h.}.$$

The value of the differences $(M_l(Y) - \bar{M}_{hl}(Y))$ is never negative, while the differences $\left(M_g(Y) - \bar{M}_{hl}(Y)\right)$ can assume positive and negative values. Consequently, the within parts are never negative and the between parts can assume positive and negative values as is shown in the last three rows of Table 28.

Tables 26, 27 and 28 report also the decomposition of the three point indexes by the four sources and the joint decompositions by subpopulations and souces.

Table 37 reports, for each of the four subpopulation l and for the whole country the decompositions by sources of the corresponding (three) Gini point and synthetic indexes.

Finally, Table 36 reports the relative contributions $\lambda_{hl.}$ and Table 38 reports the relative contributions $\lambda_{hl}(X_j) = [C_{hl}(X_j)/C_{hl}(Y)]$. It is worth while pointing out that there are very remarkable differences between these latter relative contributions. To understand why it is possible to have these remarkable differences, the paper shows, first of all, that $\lambda_{hl.}(X_j) = \frac{C_{hlW}(X_j)}{C_{hlW}(Y)} \cdot \frac{C_{hlW}(Y)}{C_{hl.}(Y)} + \frac{C_{hlB}(X_j)}{C_{hlB}(Y)} \cdot \frac{C_{hlB}(Y)}{C_{hl.}(Y)}$. Then, the paper obtains the following relations: $\lambda_{hl.}(X_j) \simeq \lambda_{hl}(X_j) \cdot \frac{C_{hlW}(Y)}{C_{hl.}(Y)} + \frac{C_{hlB}(X_j)}{C_{hlB}(Y)} \cdot \frac{C_{hlB}(Y)}{C_{hl.}(Y)}$. In other words the relative contributions $\lambda_{hl.}(X_j)$ is the weighted mean of the relative contributions $\lambda_{hl}(X_j)$ and the ratio $[C_{hlB}(X_j)/C_{hlB}(Y)]$. The relation (99) explains why it is possible to have remarkable differences between the relative contribution of X_j in the subpopulation l and the relative contributions of X_j to the contributions $C_{hl.}(Y)$ of the subpopulation l to the point Gini $G_h(Y)$.

We end this paper by pointing out that the contributions $C_{hlB}(Y)$ and $C_{hl.}(Y)$ can be negative. The fact that some contributions are negative may cause some difficulties in interpretation. This cannot happen for the Zenga (2007) inequality index. For more details see Zenga and Valli (2016, 2017), Zenga *et al.* (2012) and Zenga (2015).

REFERENCES

Bank of Italy (2016). *Household income and wealth in 2014*. Supplements to the Statistical Bulletin - Sample Survey.

Bonferroni C. (1930). *Elementi di statistica generale. Seeber, Firenze.*

Gini C. (1914). *Sulla misura della concentrazione e della variabilità dei caratteri*. In Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti. Anno Accademico 1913-1914, **LXXIII** - Parte Seconda.

Lorenz M. (1905). Methods of measuring the concentration of wealth. *American Statistical Association*, **9**, 209-219.

Pasquazzi L., Zenga M. (2018). Components of Gini, Bonferroni, and Zenga inequality Indexes for EU Income Data. *Journalof Official Statistics*, **34**(1), 149-180.

Zenga M.M. (2007). Inequality curve and inequality index based on the ratios between lower and upper means. *Statistica & Applicazioni*, **1**, 3-27.

Zenga M. (2013). Decomposition by sources of the Gini, Bonferroni and Zenga inequality indexes. *Statistica & Applicazioni*, **XI**(2), 133-161.

Zenga M. (2015). Joint decomposition by subpopulations and sources of the point and synthetic Zenga (2007) index I(Y). *Statistica & Applicazioni*, **XIII**(2), 163-195.

Zenga M. (2016a). Decomposition by subpopulations of the point and the synthetic Gini inequality indexes. *Statistics & Applications*, **XIV**(1), 49-82.

Zenga M. (2016b). On the decomposition by subpopulations of the point and synthetic Zenga (2007) inequality indexes. *Metron*, **74**(3), 375-405.

Zenga M., Valli I. (2016). On the decomposition by subpopulations of the point and synthetic Bonferroni inequality measures. *Statistics & Applications*, **XIV**(1), 3-28.

Zenga M., Valli I. (2017). Joint decomposition by subpopulations and sources of the point and synthetic Bonferroni inequality measures. *Statistics & Applications*, **XV**(2), 83-120.

Zenga M.M., Radaelli P., Zenga Ma. (2012). Decomposition of Zenga's inequality index by sources. *Statistica & Applicazioni,* **X**(1), 3-31.