Statistica & Applicazioni - Vol. XV, n. 2, 2017, pp. 83-120

JOINT DECOMPOSITION BY SUBPOPULATIONS AND SOURCES OF THE POINT AND SYNTHETIC BONFERRONI INEQUALITY MEASURES

Michele Zenga*
Igor Valli*

ABSTRACT

The total income Y is the sum of c sources $X_j : Y = X_1 + \ldots + X_c$. The N units of the population are partitioned in k different subpopulations. In the frequency distribution framework the Bonferroni (1930) point inequality index is given by $V_h(Y) = [M(Y) - \bar{M}_{h.}(Y)]/M(Y)$, $M(Y)$ and $\bar{M}_{h.}(Y)$ are the mean and the lower mean of Y. The synthetic Bonferroni index $V(Y)$ is the mean $M(\cdot)$ of $V_h(Y) : V(Y) = M[V_h(Y)]$. Zenga and Valli (2016) have decomposed $V_h(Y)$ in a weighted mean of the $k \times k$ relative differences $\left[M_g(Y) - \bar{M}_{h\ell}(Y)\right]/M(Y)$; $M_g(Y)$ is the mean of subpopulation g and $\bar{M}_{h\ell}(Y)$ is the lower mean of subpopulation ℓ, the weights are the product $p(\ell|h) \cdot n_{.g}/N$, where $n_{.g}/N$ is the relative frequency of the subpopulation g and $p(\ell|h)$ is the relative frequency of the subpopulation ℓ in the lower group. From this $k \times k$ additive decomposition, the Authors obtained the decomposition of $V_h(Y)$ into the within and the between components as well as the decomposition of $V_h(Y)$ into the sum of the subpopulation contributions: $V_{h\ell.}(Y) = \frac{1}{M(Y)} \cdot [M(Y) - \bar{M}_{h\ell}(Y)] \cdot p(\ell|h)$. Using the relations $\bar{M}_{hg}(Y) = \sum_{j=1}^{c} \bar{M}_{hg}(X_j)$ and $M(Y) = \sum_{j=1}^{c} M(X_j)$ in the $k \times k$ decomposition of $V_h(Y)$, the present paper obtains the following $k \times k \times c$ joint decomposition by subpopulations and sources: $V_h(Y) = \sum_\ell \sum_g \sum_j V_{h\ell g}(X_j)$; $V_{h\ell g}(X_j) = \frac{M_g(X_j) - \bar{M}_{h\ell}(X_j)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.g}}{N}$. The joint decomposition of $V(Y)$ is obtained putting the joint decomposition of $V_h(Y)$ in $M[V_h(Y)]$. The decompositions proposed in this paper are applied to the net disposable income Y of the 8156 Italian households partitioned in three macro-regions, supplied by the 2014 Central Bank of Italy sample survey. The total income Y is the sum of four sources.

Keywords: *Point and Synthetic Bonferroni (1930) Inequality Indexes, Decomposition by Subpopulations, Decomposition by Sources, Joint Decomposition by Subpopulations and Sources, Income Inequality.*

1. INTRODUCTION

The total income Y is the sum of c sources $X_j : Y = X_1 + \ldots + X_j + \ldots + X_c$. The N units of the population are partitioned in k different subpopulations. In the case of N values of Y arranged in non decreasing order $\left\{0 \le y_{(1)} \le \ldots \le y_{(i)} \le \ldots \le y_{(N)} > 0\right\}$, Zenga (2013) using the ''two-step'' approach obtained the decomposition by sources of the points and the synthetics Gini (1914), Bonferroni (1930) and Zenga (2007) inequality indexes. In the case of frequency distribution framework, $\{0 \le y_1 < \ldots < y_h < \ldots < y_r\}$ denotes the set of

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

the r distinct values assumed by Y over the k subpopulations and $\left\{n_{1.}, \ldots, n_{h.}, \ldots n_{r.}; \sum_{h=1}^{r} n_{h.} = N\right\}$ are the corresponding frequencies. In this latter case the Bonferroni point index $V_h(Y)$ is given by $V_h(Y) = \frac{M(Y)-\bar{M}_{h.}(Y)}{M(Y)}$; where $M(Y)$ is the mean of Y and $\bar{M}_{h.}(Y)$ is the lower mean (the mean of Y computed on the $P_{h.} = \sum_{t=1}^{h} n_{t.}$ units of the lower group $\{Y \leq y_h\}$). The Bonferroni synthetic index $V(Y)$ is the weighted mean of $V_h(Y) : V(Y) = \sum_{h=1}^{r} V_h(Y) \cdot \frac{n_{h.}}{N}$. Recently, Zenga and Valli (2016) have decomposed $V_h(Y)$ in a weighted mean of the $k \times k$ relative differences $\frac{M_g(Y)-\bar{M}_{h\ell}(Y)}{M(Y)}$: $M_g(Y)$ in the mean of subpopulation g and $\bar{M}_{h\ell}(Y)$ is the lower mean of subpopulation ℓ; the weights are the products $p(\ell|h) \cdot \frac{n_{.g}}{N}$, $\frac{n_{.g}}{N}$ is the relative frequency of subpopulation g and $p(\ell|h)$ is the relative frequency of subpopulation ℓ in the lower group $\{Y \leq y_h\}$. From this $k \times k$ additive decomposition of $V_h(Y)$ the Authors obtain the decomposition of $V_h(Y)$:

1. in the sum of the contributions $V_{h\ell.}(Y)$ of each subpopulation:

$$V_{h\ell.}(Y) = \frac{M(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot p(\ell|h), \quad \ell = 1, \ldots, k;$$

2. in the sum of the within and the between components.

The present paper obtains (first step) the following $k \times k \times c$ joint decomposition by subpopulations and sources of the Bonferroni point index $V_h(Y) = \sum_{\ell=1}^{k} \sum_{g=1}^{k} \sum_{j=1}^{c} V_{h\ell g}(X_j)$, where $V_{h\ell g}(X_j) = \frac{M_g(X_j)-\bar{M}_{h\ell}(X_j)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.g}}{N}$. The joint decomposition of $V(Y)$ is obtained (second step) putting this latter decomposition of $V_h(Y)$ in $V(Y) = \sum_{h=1}^{r} V_h(Y) \cdot \frac{n_{h.}}{N}$. The rest of the paper is organized as follows. In the next section some definitions and notation are given for the case of frequency distribution framework. In particular this section provides the definitions: of the lower mean and of the mean of the variate Y in the whole population, and of the point $V_h(Y)$ and the synthetic Bonferroni indexes. Moreover, this section reports the properties of the arithmetic mean that play an important role in the decomposition by sources, by subpopulations and in the joint decomposition. Section 3 illustrates the ‘‘two-step’’ approach for the additive decomposition by sources of the point $V_h(Y)$ and the synthetic Bonferroni indexes. The bivariate distribution of the N units according to the 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 the definitions of the mean and the lower mean. Finally, this section illustrates the basic $k \times k$ additive decomposition by subpopulations of $V_h(Y)$ and the $k \times k$ additive decomposition of $V(Y)$. Section 5 obtains the basic $k \times k \times c$ joint decomposition by subpopulations and sources of the point Bonferroni index as well as the $k \times k \times c$ additive decomposition by subpopulations and sources of the synthetic Bonferroni index $V(Y)$. In particular, Subsection 5.1 introduces, for each subpopulation g, the means $M_g(X_j)$ and the lower means $\bar{M}_{hg}(X_j)$ of the c sources.

Using the relations $\bar{M}_{hg}(Y) = \sum_{j=1}^{c} \bar{M}_{hg}(X_j)$ and $M_g(Y) = \sum_{j=1}^{c} M_g(X_j)$ in the $k \times k$ decomposition (by subpopulations) of $V_h(Y)$, Section 5.2 obtains (first step) the following $k \times k \times c$ additive decomposition of $V_h(Y) : V_h(Y) = = \sum_{\ell=1}^{k} \sum_{g=1}^{k} \sum_{j=1}^{c} V_{h\ell g}(X_j)$, where $V_{h\ell g}(X_j) = \frac{M_g(X_j) - \bar{M}_{h\ell}(X_j)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.g}}{N}$. In the usual way (second step) this section provides the joint decomposition of $V(Y)$. Section 6 illustrates for each subpopulation, the decomposition by sources of the point and the synthetic Bonferroni inequality indexes. Section 7 provides an application of the decompositions proposed in this paper to the net disposable income Y of the $N = 8156$ Italian households partitioned in three macro-regions (North, Center and South) supplied by the 2014 Central Bank of Italy sample survey. The total income Y is the sum of the: payroll income (X_1), pensions and net transfers (X_2), net self employment (X_3), property income (X_4). Conclusions and final remarks end the paper.

2. DEFINITIONS AND NOTATION

Let, $X_1, \ldots, X_j, \ldots, X_c$ be c non-negative variates (incomesources) 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 r distinct values assumed by the variate Y and $\{n_{1.}, \ldots, n_{h.}, \ldots, n_{r.}\}$ be the corresponding frequencies: $\sum_{h=1}^{r} n_{h.} = N$.

At each y_h we can consider a lower group $\{Y \leq y_h\}$, including the first

$$P_{h.} = \sum_{t=1}^{h} n_{t.} \text{ units, with cumulative income } Q_{h.}(Y) = \sum_{t=1}^{h} y_t \cdot n_{t.} \tag{1}$$

Let,

$$\bar{M}_{h.}(Y) = \frac{Q_{h.}(Y)}{P_{h.}} \qquad h = 1, \ldots, r, \tag{2}$$

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

$$M(Y) = \frac{Q_{r.}(Y)}{N} \tag{3}$$

be the arithmetic mean of the overall population.

In the frequency distribution framework (see Zenga and Valli, 2016), the Bonferroni (1930) point $V_h(Y)$ and synthetic $V(Y)$ indexes are given by:

$$V_h(Y) = \frac{M(Y) - \bar{M}_{h.}(Y)}{M(Y)}, \qquad h = 1, \ldots, r; \tag{4}$$

$$V(Y) = \sum_{h=1}^{r} V_h(Y) \cdot \frac{n_{h.}}{N}. \tag{5}$$

The following three properties of the arithmetic mean M, play an important role in the decompositions by sources, by subpopulations and in the joint decomposition by subpopulations and sources.

Property I. *Let,* $X_1, \ldots, X_j, \ldots, X_c$ *be* c *variates observable on each of the* N *units of a population and* $Y = \sum_{j=1}^{c} X_j$ *be their sum. Then:*

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

Property II. *The* N *units of a population are split in* k *different subpopulations and* $\left\{n_{.1}, \ldots n_{.g}, \ldots, n_{.k} : \sum_{g=1}^{k} n_{.g} = N\right\}$ *are their sizes. Then:*

$$M = \sum_{g=1}^{k} M_g \cdot \frac{n_{.g}}{N},$$

where M is the mean of the whole population and M_g is the mean of subpopulation g, $g = 1, \ldots k$.

Property III. *Let* $\left\{(v_i, n_i) : \ i = 1, \ldots, a; \ \sum_{i=1}^{a} n_i = n\right\}$ *be the frequency distribution of the variate* V *in the population* A *and* $\left\{(z_j, m_j) : \ j = 1, \ldots, b; \ \sum_{j=1}^{b} m_j = m\right\}$ *be the frequency distribution of the variate* Z *in population* $B.M(V) = \frac{1}{n}\sum_{i=1}^{a} v_i \cdot n_i$ *and* $M(Z) = \frac{1}{m}\sum_{j=1}^{b} z_j \cdot m_j$ *are the means of* V *and* Z*. We want evaluate the mean of the* $n \times m$ *differences* $(V - Z)$ *obtained comparing each units in* A *with each units in* B*. Then,*

$$\begin{aligned} M(V - Z) &= \sum_{i=1}^{a} \sum_{j=1}^{b} (v_i - z_j) \frac{n_i m_j}{n \cdot m} = \sum_{j=1}^{b} \frac{m_j}{m} \sum_{i=1}^{a} v_i \frac{n_i}{n} - \sum_{i=1}^{a} \frac{n_i}{n} \sum_{j=1}^{b} z_j \frac{m_j}{m} = \\ &= M(V) - M(Z). \end{aligned}$$

3. Decomposition by income sources

Let, $Q_{h.}(X_j)$, $j = 1, \ldots, c$, $h = 1, \ldots, r$, be the sum of the values of X_j observable in the lower group: $Q_{r.}(X_j) = T(X_j)$ is the sum of all the N values of X_j. Let,

$$\bar{M}_{h.}\left(X_j\right) = \frac{Q_{h.}\left(X_j\right)}{P_{h.}}, \tag{6}$$

be the mean of X_j in the lower group, and

$$M(X_j) = \frac{Q_{r.}(X_j)}{N} \tag{7}$$

be the mean of X_j in the whole population. Then, from Property I:

$$\bar{M}_{h.}(Y) = \sum_{j=1}^{c} \bar{M}_{h.}(X_j) \quad \text{and} \quad M(Y) = \sum_{j=1}^{c} M(X_j). \tag{8}$$

Using relations (8) in the numerator of (4), the following decomposition by sources of $V_h(Y)$ is obtained:

$$V_h(Y) = \frac{\sum_{j=1}^{c} M(X_j) - \sum_{j=1}^{c} \bar{M}_{h.}(X_j)}{M(Y)} = \sum_{j=1}^{c} V_h(X_j), \tag{9}$$

where,

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

is the contribution of X_j to $V_h(Y)$.

Finally, putting the decomposition (9) in (5), the decomposition by sources of $V(Y)$ is obtained:

$$V(Y) = \sum_{h=1}^{r} \left[\sum_{j=1}^{c} V_h(X_j) \right] \cdot \frac{n_{h.}}{N} = \sum_{j=1}^{c} \sum_{h=1}^{r} V_h(X_j) \cdot \frac{n_{h.}}{n} = \sum_{j=1}^{c} V(X_j), \tag{11}$$

where,

$$V(X_j) = \sum_{h=1}^{r} V_h(X_j) \cdot \frac{n_{h.}}{N} \tag{12}$$

is the contribution of X_j to the syntetic index $V(Y)$. Note that $V(X_j)$ is the arithmetic mean of the contributions of X_j to the point indexes $V_h(Y)$.

4. DECOMPOSITION BY SUBPOPULATIONS

In the case of the decomposition by subpopulations we need the bivariate distribution of the N units according to k disjoint subpopulations and the r distinct values of Y. This distribution is reported in Table 1, where n_{hg} denotes the frequency of the value y_h in the subpopulation g and $n_{.g}$ is the size of the subpopulation g.

TABLE 1. - *Bivariate distribution of the N units according to the r distinct values of Y and the k subpopulations*

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

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

$$P_{hg} = \sum_{t=1}^{h} n_{tg}, \quad \text{and} \quad Q_{hg}(Y) = \sum_{t=1}^{h} y_t \cdot n_{tg} \; . \tag{13}$$

Moreover:

$$T_g(Y) = Q_{rg}(Y) \quad \text{and} \quad M_g(Y) = T_g(Y)/n_{.g} \; . \tag{14}$$

In addition, we define for the subpopulation g the lower mean $\bar{M}_{hg}(Y)$ as follows:

$$\bar{M}_{hg}(Y) = \begin{cases} y_{o(g)} & \text{for } h < o(g) \\ Q_{hg}(Y)/P_{hg} & \text{for } h \geq o(g) \end{cases} \tag{15}$$

where,

$$o(g) = \min h : n_{hg} > \; 0. \tag{16}$$

The decomposition by subpopulations of $V_h(Y)$ (see Zenga and Valli, 2016) is obtained starting from the decomposition of the difference

$$[M(Y) - \bar{M}_{h.}(Y)].$$

(i) By the use of the Property II,

$$M(Y) = \sum_{g=1}^{k} M_g(Y) \cdot \frac{n_{.g}}{N}, \quad \text{and} \tag{17}$$

$$\bar{M}_{h}.(Y) = \sum_{\ell=1}^{k} \bar{M}_{h\ell}(Y) \cdot p(\ell|h) \quad \text{where;} \tag{18}$$

$$p(\ell|h) = \frac{P_{h\ell}}{P_{h.}} \qquad h = 1, \ldots, r \tag{19}$$

is the relative frequency of the subpopulation ℓ in the lower group. Now, by the use of Property III, we obtain the following decomposition of $[M(Y) - \bar{M}_{h.}(Y)]$:

$$[M(Y) - \bar{M}_{h.}(Y)] = \sum_{\ell=1}^{k} \sum_{g=1}^{k} \left[M_g(Y) - \bar{M}_{h\ell}(Y)\right] \cdot p(\ell|h) \cdot \frac{n_{.g}}{N} \; . \tag{20}$$

(ii) Dividing both sides of (20) by $M(Y)$, Zenga and Valli (2016) have obtained the following *basic* $k \times k$ additive decomposition of $V_h(Y)$:

$$V_h(Y) = \frac{M(Y) - \bar{M}_h.(Y)}{M(Y)} = \sum_{\ell=1}^{k} \sum_{g=1}^{k} \frac{M_g(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} p(\ell|h) \frac{n_{.g}}{N}$$

$$= \sum_{\ell=1}^{k} \sum_{g=1}^{k} V_{h\ell g}(Y), \tag{21}$$

where

$$V_{h\ell g}(Y) = \frac{M_g(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.g}}{N} \tag{22}$$

is the contribution to the point inequality index $V_h(Y)$ that derives from the comparison of the lower mean $\bar{M}_{h\ell}(Y)$ with the mean $M_g(Y)$.

It is worth to remark that, starting from the $k \times k$ decomposition (21), it is possible to obtain many other additive decompositions of $V_h(Y)$. Thus, after some steps we obtain:

$$V_{h\ell.}(Y) = \sum_{g=1}^{k} V_{h\ell g}(Y) = \frac{M(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot p(\ell|h). \tag{23}$$

Putting (23) in (21) gives,

$$V_h(Y) = \sum_{\ell=1}^{k} V_{h\ell.}(Y) = \sum_{\ell=1}^{k} \frac{M(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot p(\ell|h). \tag{24}$$

Formula (24) informs that, the point index $V_h(Y)$ is the weighted mean of the k relative variations

$$\left[\frac{M(Y) - \bar{M}_{h\ell}(Y)}{M(Y)}\right]$$

with weights $p(\ell|h)$. Thus, $V_{h\ell.}(Y)$ can be interpreted as the contribution of the subpopulation ℓ to the point inequality index $V_h(Y)$. It is worth to remark that the other main decompositions proposals do not derive the contribution of each subpopulation to the (synthetic) Bonferroni index.

Now, $V_{h\ell.}(Y)$ can be split into a within and a between components.

$$\begin{aligned} V_{h\ell.}(Y) &= \sum_{g=1}^{k} V_{h\ell g}(Y) = V_{h\ell\ell}(Y) + \sum_{(g:g\neq\ell)} V_{h\ell g}(Y) \\ &= V_{h\ell W}(Y) + V_{h\ell B}(Y). \end{aligned} \tag{25}$$

Consequently, the within and the between component of the point index $V_h(Y)$ are given by:

$$V_h(Y) = \sum_{\ell=1}^{k} V_{h\ell.}(Y) = \sum_{\ell=1}^{k} V_{h\ell W}(Y) + \sum_{\ell=1}^{k} V_{h\ell B}(Y) = V_{h.W}(Y) + V_{h.B}(Y), \tag{26}$$

where $V_{h.W}(Y) = \sum_{\ell=1}^{k} V_{h\ell W}(Y)$ can be interpreted as the within contribution of all subpopulations to $V_h(Y)$, and $V_{h.B}(Y) = \sum_{\ell=1}^{k} V_{h\ell B}(Y)$ can be interpreted as the between contribution of all the subpopulations to $V_h(Y)$.

Finally, putting the decomposition (26) of $V_h(Y)$ in (5), the following decomposition of the synthetic index is obtained:

$$V(Y) = \sum_{\ell=1}^{k} \sum_{g=1}^{k} \left[\sum_{h=1}^{r} V_{h\ell g}(Y) \cdot \frac{n_{h.}}{N} \right] = \sum_{\ell=1}^{k} \sum_{g=1}^{k} V_{.\ell g}(Y). \tag{27}$$

In (27) $V_{.\ell g}(Y)$ is the weighted mean of the r contributions $V_{h\ell g}(Y)$. Now, from (27) we obtain:

$$V(Y) = \sum_{\ell=1}^{k} \left[V_{.\ell\ell}(Y) + \sum_{(g:g\neq\ell)} V_{.lg}(Y) \right] = \sum_{\ell=1}^{k} V_{.\ell.}(Y) \tag{28}$$

where

$$V_{.\ell.}(Y) = \sum_{g=1}^{k} V_{.\ell g}(Y) \tag{29}$$

is the contribution of subpopulation ℓ to the synthetic index $V(Y)$. Moreover, from the relation (28) we have:

$$V(Y) = \sum_{\ell=1}^{k} [V_{.\ell W}(Y) + V_{.\ell B}(Y)] = V_{..W}(Y) + V_{..B}(Y). \tag{30}$$

In (30):

$$V_{.\ell W}(Y) = V_{.\ell\ell}(Y) \quad \text{and} \quad V_{.\ell B}(Y) = \sum_{(g:g\neq\ell)} V_{.\ell g}(Y) \tag{31}$$

are the within and the between contributions of the subpopulation ℓ to $V(Y)$;

$$V_{..W} = \sum_{\ell=1}^{k} V_{.\ell W}(Y) \quad \text{and} \quad V_{..B}(Y) = \sum_{\ell=1}^{k} V_{.\ell B}(Y) \tag{32}$$

can be interpreted respectively as the within and the between components of the synthetic index $V(Y)$.

It is worth to remark that in

$$V_{..W}(Y) = \sum_{\ell=1}^{k} \left[\sum_{h=1}^{r} \frac{M_\ell(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.\ell}}{N} \cdot \frac{n_{h.}}{N} \right], \tag{33}$$

only comparisons between means and lower means of the same subpopulation are involved, while in

$$V_{..B}(Y) = \sum_{\ell=1}^{k} \sum_{(g:g\neq\ell)} \left[\sum_{h=1}^{r} \frac{M_g(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.g}}{N} \cdot \frac{n_{h.}}{N} \right] \tag{34}$$

only comparisons between means and lower means of different subpopulations are involved.

4.1 *A more direct way to obtain the contribution of each subpopulation to the point Bonferroni index*

It is worth to remark that by the use of relation (18) in (4) we can obtain, in a more direct way, the decomposition (24).

$$
\begin{aligned}
V_h(Y) &= \frac{M(Y) - \bar{M}_{h.}(Y)}{M(Y)} = \frac{M(Y) - \sum_{\ell=1}^{k} \bar{M}_{h\ell}(Y) p(\ell|h)}{M(Y)} \\
&= \frac{M(Y) \sum_{\ell=1}^{k} p(\ell|h) - \sum_{\ell=1}^{k} \bar{M}_{h\ell}(Y) p(\ell|h)}{M(Y)} = \sum_{\ell=1}^{k} \frac{M(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} p(\ell|h).
\end{aligned}
$$

Now, by the use of relation (17) in the numerator of $V_{h\ell.}(Y)$ the within and the between components of $V_{h\ell.}(Y)$ are obtained.

$$
\begin{aligned}
V_{h\ell.}(Y) &= \frac{\sum_{g=1}^{k} M_g(Y) \cdot \frac{n_{.g}}{N} - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot p(\ell|h) \\
&= \frac{\sum_{g=1}^{k} M_g(Y) \cdot \frac{n_{.g}}{N} - \bar{M}_{h\ell}(Y) \cdot \sum_{g=1}^{k} \frac{n_{.g}}{N}}{M(Y)} \cdot p(\ell|h) \\
&= \sum_{g=1}^{k} \frac{M_g(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot \frac{n_{.g}}{N} \cdot p(\ell|h) \\
&= \frac{M_\ell(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot \frac{n_{.\ell}}{N} \cdot p(\ell|h) + \sum_{g:g\neq\ell} \frac{M_g(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot \frac{n_{.g}}{N} \cdot p(\ell|h).
\end{aligned}
$$

Finally, from this latter decomposition of $V_{h\ell.}(Y)$ we can obtain the within and the between parts of the point index $V_h(Y)$.

5. JOINT DECOMPOSITION OF $V_h(Y)$ AND $V(Y)$ BY SUBPOPULATIONS AND SOURCES

To get the joint decomposition we need, for each subpopulation g, the lower means $\bar{M}_{hg}(X_j)$ and the means $M_g(X_j)$ of the c sources. Subsection 5.1 introduces the lower means $\bar{M}_{hg}(X_j)$ and the means $M_g(X_j)$ of the c sources and their relationship with the corresponding lower means and means of the sum Y. Then, in Subsection 5.2 the joint decomposition by subpopulations and sources is achieved.

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

Let,

$$S_{hg}(Y) = y_h \cdot n_{hg}, \qquad h = 1, \ldots, r; \qquad g = 1, \ldots, k. \tag{35}$$

Thus, from (13),

$$Q_{hg}(Y) = \sum_{t=1}^{h} y_t \cdot n_{tg} = \sum_{t=1}^{h} S_{tg}(Y). \tag{36}$$

In the case $n_{hg} > 0$, let us denote with x_{hgjd}, $d = 1, \ldots, n_{hg}$, 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, $\sum_{j=1}^{c} x_{hgjd} = y_h$, $\forall d = 1, \ldots, n_{hg}$, $\forall h = 1, \ldots, r$, $\forall g = 1, \ldots, k$.

Let:

$$S_{hg}(X_j) = \begin{cases} 0 & \text{for } n_{hg} = 0 \\ \sum_{d=1}^{n_{hg}} x_{hgjd} & \text{for } n_{hg} > 0, \end{cases} \tag{37}$$

$$M_{hg}(X_j) = \frac{S_{hg}(X_j)}{n_{hg}} \qquad \text{for } n_{hg} > 0, \tag{38}$$

$$M_{hg}(Y) = \frac{S_{hg}(Y)}{n_{hg}} = y_h \qquad \text{for } n_{hg} > 0. \tag{39}$$

Note that:

$$\sum_{d=1}^{n_{hg}} \sum_{j=1}^{c} x_{hgjd} = \sum_{d=1}^{n_{hg}} y_h = y_h \cdot n_{hg} = S_{hg}(Y),$$

$$\sum_{j=1}^{c} \sum_{d=1}^{n_{hg}} x_{hgjd} = \sum_{j=1}^{c} S_{hg}(X_j) = S_{hg}(Y), \tag{40}$$

$$\sum_{j=1}^{c} M_{hg}(X_j) \cdot n_{hg} = M_{hg}(Y) \cdot n_{hg}. \tag{41}$$

Let:

$$Q_{hg}(X_j) = \sum_{t=1}^{h} S_{tg}(X_j), \tag{42}$$

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:

$$T_g(X_j) = Q_{rg}(X_j) = \sum_{h=1}^{r} S_{hg}(X_j),$$

be the sum of the $n_{.g}$ values of X_j of the subpopulation g, and

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

Now, from (42), (40) and (36):

$$\sum_{j=1}^{c} Q_{hg}(X_j) = \sum_{j=1}^{c} \sum_{t=1}^{h} S_{tg}(X_j) = \sum_{t=1}^{h} \sum_{j=1}^{c} S_{tg}(X_j) = \sum_{t=1}^{h} S_{tg}(Y) = Q_{hg}(Y). \tag{43}$$

In coherence with (15) let:

$$\bar{M}_{hg}(X_j) = \begin{cases} M_{o(g)g}(X_j) & \text{for } h < o(g) \\ \frac{Q_{hg}(X_j)}{P_{hg}} & \text{for } h \geq o(g). \end{cases} \tag{44}$$

Finally, for $h = 1, \ldots, r$:

$$\sum_{j=1}^{c} \bar{M}_{hg}(X_j) = \begin{cases} \sum_{j=1}^{c} M_{o(g)g}(X_j) = M_{o(g)g}(Y) & \text{for } h < o(g) \\ \sum_{j=1}^{c} \frac{Q_{hg}(X_j)}{P_{hg}} = \frac{Q_{hg}(Y)}{P_{hg}} & \text{for } h \geq o(g) \end{cases}$$
$$= \bar{M}_{hg}(Y), \tag{45}$$

and

$$M_g(Y) = \sum_{j=1}^{c} M_g(X_j). \tag{46}$$

5.2 *From the decomposition by subpopulations to the joint decomposition*

By the use of the relations (45) and (46) in the decomposition by subpopulations (21) of $V_h(Y)$, the *basic* $k \times k \times c$ joint decomposition by subpopulations and sources of $V_h(Y)$ is obtained:

$$V_h(Y) = \sum_{\ell=1}^{k} \sum_{g=1}^{k} \sum_{j=1}^{c} V_{h\ell g}(X_j), \tag{47}$$

where

$$V_{h\ell g}(X_j) = \left[\frac{M_g(X_j) - \bar{M}_{h\ell}(X_j)}{M(Y)} \right] \cdot p(\ell|h) \cdot \frac{n_{.g}}{N} \tag{48}$$

is the contribution to $V_h(Y)$ that derives from the comparison of the lower mean $\bar{M}_{h\ell}(X_j)$ w.r.t the mean $M_g(X_j)$. Note that:

$$\sum_{j=1}^{c} V_{h\ell g}(X_j) = V_{h\ell g}(Y). \tag{49}$$

In other words the contribution $V_{h\ell g}(Y)$ is decomposed additively by the contributions $V_{h\ell g}(X_j)$ of each component X_j of the sum Y.

Now, after some steps we obtain:

$$V_{h\ell .}(Y) = \sum_{j=1}^{c} V_{h\ell .}(X_j), \tag{50}$$

where

$$V_{h\ell .}(X_j) = \frac{M(X_j) - \bar{M}_{h\ell}(X_j)}{M(Y)} \cdot p(\ell|h) \tag{51}$$

is the ‘‘additive’’ contribution of X_j to $V_{h\ell .}(Y)$. Moreover,

$$V_{h\ell .}(X_j) = V_{h\ell W}(X_j) + V_{h\ell B}(X_j), \tag{52}$$

where

$$V_{h\ell W}(X_j) = V_{h\ell\ell}(X_j), \tag{53}$$

and

$$V_{h\ell B}(X_j) = \sum_{(g:g\neq\ell)} V_{h\ell g}(X_j), \tag{54}$$

are the within and the between parts of $V_{h\ell .}(X_j)$. Now from (50) and (52) we have:

$$V_{h\ell .}(Y) = \sum_{j=1}^{c} \left[V_{h\ell W}(X_j) + V_{h\ell B}(X_j)\right]. \tag{55}$$

Finally, putting (55) in (26) the following ‘‘joint decompositions’’ of the point index $V_h(Y)$ are obtained:

$$V_h(Y) = \sum_{\ell=1}^{k} \sum_{j=1}^{c} \left[V_{h\ell W}(X_j) + V_{h\ell B}(X_j)\right]$$

$$V_h(Y) = \sum_{j=1}^{c} \sum_{\ell=1}^{k} \left[V_{h\ell W}(X_j) + V_{h\ell B}(X_j)\right] \tag{56}$$

$$V_h(Y) = \sum_{j=1}^{c} \left[V_{h.W}(X_j) + V_{h.B}(X_j)\right], \tag{57}$$

where

$$V_{h.W}(X_j) = \sum_{\ell=1}^{k} V_{h\ell W}(X_j) \quad \text{and} \quad V_{h.B}(X_j) = \sum_{\ell=1}^{k} V_{h\ell B}(X_j) \tag{58}$$

are the within and the between contributions of X_j to $V_h(Y)$. Finally, putting the *basic* joint decomposition (47) of $V_h(Y)$ in (4), the $k \times k \times c$ joint decomposition (by subpopulations and sources) of $V(Y)$ is obtained:

$$V(Y) = \sum_{\ell=1}^{k} \sum_{g=1}^{k} \sum_{j=1}^{c} \sum_{h=1}^{r} V_{h\ell g}(X_j) \frac{n_{h.}}{N} = \sum_{\ell=1}^{k} \sum_{g=1}^{k} \sum_{j=1}^{c} V_{.\ell g}(X_j), \tag{59}$$

where

$$V_{.\ell g}(X_j) = \sum_{h=1}^{r} \frac{M_g(X_j) - \bar{M}_{h\ell}(X_j)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.g}}{N} \cdot \frac{n_{h.}}{N} \tag{60}$$

is the weighted arithmetic mean of $V_{h\ell g}(X_j)$ with weights $\frac{n_{h.}}{N}$.

Starting from the $k \times k \times c$ joint decomposition (59) of $V(Y)$, it is possible to derive many other decompositions. In Section 7 some of these decompositions will be illustrated.

5.3 *From the decomposition by sources to the joint decomposition*

We have to remark that the joint decomposition (47) of $V_h(Y)$ can also be obtained starting from the decomposition by sources (9) of $V_h(Y)$:

$$V_h(Y) = \sum_{j=1}^{c} \frac{M(X_j) - \bar{M}_{h.}(X_j)}{M(Y)} = \sum_{j=1}^{c} V_h(X_j),$$

where

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

By the use of Property II (Section 2) the means of the numerator of (10) are given by:

$$M(X_j) = \sum_{g=1}^{k} M_g(X_j) \cdot \frac{n_{.g}}{N} \quad \text{and} \quad \bar{M}_{h.}(X_j) = \sum_{\ell=1}^{k} \bar{M}_{h\ell}(X_j) \cdot p(\ell|h).$$

Now, using (first step) these latter relations in the numerator of (10) we achieve the following $k \times k$ representation of $V_h(X_j)$:

$$V_h(X_j) = \sum_{\ell=1}^{k} \sum_{g=1}^{k} \frac{M_g(X_j) - \bar{M}_{h\ell}(X_j)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.g}}{N}.$$

Finally, putting in (9) this latter representation of $V_h(X_j)$, the result is reached:

$$\begin{aligned} V_h(Y) &= \sum_{j=1}^{c} \sum_{\ell=1}^{k} \sum_{g=1}^{k} \left[\frac{M_g(X_j) - \bar{M}_{h\ell}(X_j)}{M(Y)} \right] \cdot p(\ell|h) \cdot \frac{n_{.g}}{N} \\ &= \sum_{\ell=1}^{k} \sum_{g=1}^{k} \sum_{j=1}^{c} V_{h\ell g}(X_j), \end{aligned} \tag{47}$$

where

$$V_{h\ell g}(X_j) = \frac{M_g(X_j) - \bar{M}_{h\ell}(X_j)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.g}}{N} \tag{48}$$

5.4 *Relative contributions of the component X_j to the point $V_h(Y)$ and to the synthetic $V(Y)$ indexes*

In the context of the joint decomposition, it is useful to denote:

a) the sum of the $k \times k$ contributions $V_{h\ell g}(X_j)$ with

$$V_{h..}(X_j) = \sum_{\ell=1}^{k} \sum_{g=1}^{k} V_{h\ell g}(X_j), \tag{61}$$

obviously $V_{h..}(X_j)$ is equal to $V_h(X_j)$ introduced in Section 3;

$b)$ the contribution of X_j to $V(Y)$ with

$$V_{...}(X_j) = \sum_{h=1}^{r} V_{h..}(X_j) \cdot \frac{n_{h.}}{N}, \tag{62}$$

and it is equal to the contribution $V(X_j)$ introduced in Section 3.

If $V_h(Y) > 0$, we may define the relative contribution of X_j to the point index $V_h(Y)$ by:

$$\nu_{h..}(X_j) = \frac{V_{h..}(X_j)}{V_h(Y)} = \frac{M(X_j) - \bar{M}_{h.}(X_j)}{M(Y) - \bar{M}_{h.}(Y)}. \tag{63}$$

Obviously $\sum_{j=1}^{c} \nu_{h..}(X_j) = 1$.

The relative contribution of X_j to $V(Y)$ is given by:

$$\nu_{...}(X_j) = \frac{V_{...}(X_j)}{V(Y)} = \sum_{h=1}^{r} \nu_{h..}(X_j) \cdot \frac{V_h(Y) \cdot n_{h.}}{V(Y) \cdot N} \tag{64}$$

It has been observed in Zenga, Radaelli and Zenga (2012) that, the relative contributions to the point and the synthetic indexes should be compared with the shares $\gamma_{..}(X_j) = \frac{M(X_j)}{M(Y)}$ in order to explain how income sources X_j affect total income inequality. In this terms, if $\nu_{...}(X_j)$ is less (greater) than $\gamma_{..}(X_j)$, income sources X_j reduces (augments) inequality of total income distribution. More details on this point can be found in Zenga (2013), too.

6. DECOMPOSITION BY SOURCES FOR EACH SUBPOPULATION

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

$$V_{.\ell}(Y) = \sum_{h=1}^{r} V_{h\ell}(Y) \cdot \frac{n_{h\ell}}{n_{.\ell}} = \sum_{h=1}^{r} \frac{M_\ell(Y) - \bar{M}_{h\ell}(Y)}{M_\ell(Y)} \cdot \frac{n_{h\ell}}{n_{.\ell}}, \tag{65}$$

where

$$V_{h\ell}(Y) = \frac{M_\ell(Y) - \bar{M}_{h\ell}(Y)}{M_\ell(Y)} \tag{66}$$

is the point inequality index of the subpopulation ℓ. It can be useful to decompose by sources the point $V_{h\ell}(Y)$ and the synthetic $V_{.\ell}(Y)$ inequality indexes. By the use of Property I in the numerator of (66) we have:

$$V_{h\ell}(Y) = \sum_{j=1}^{c} \frac{M_\ell(X_j) - \bar{M}_{h\ell}(X_j)}{M_\ell(Y)} = \sum_{j=1}^{c} V_{h\ell}(X_j) \tag{67}$$

where

$$V_{h\ell}(X_j) = \frac{M_\ell(X_j) - \bar{M}_{h\ell}(X_j)}{M_\ell(Y)} \tag{68}$$

is the contribution of X_j to $V_{h\ell}(Y)$. Finally, putting the decomposition (67) in (65), the decomposition by sources of $V_{.\ell}(Y)$ is obtained:

$$V_{.\ell}(Y) = \sum_{h=1}^{r} \left[\sum_{j=1}^{c} V_{h\ell}(X_j) \right] \cdot \frac{n_{h\ell}}{n_{.\ell}} = \sum_{j=1}^{c} V_{.\ell}(X_j), \quad \text{where} \tag{69}$$

$$V_{.\ell}(X_j) = \sum_{h=1}^{r} V_{h\ell}(X_j) \cdot \frac{n_{h\ell}}{n_{.\ell}}. \tag{70}$$

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. The $N = 8156$ households have been partitioned according to their residence area: North (1), Center (2) and South with Islands (3). In all computations that follow we consider the weights $\omega_i > 0$ ($i = 1, 2, \dots, 8156$; $W = \sum \omega_i = 8156$) supplied by the Central Bank of Italy for each household; these weights are defined as the inverse of household's probability of the 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 semplicity's sake we will continue the use of the previous sections. Thus, to denote the sum of the weights of the $n_{h\ell}$ households of the subpopulation ℓ with total income $Y = y_h$ we will use $n_{h\ell}$ instead of $w_{h\ell}$. Note that the frequency distribution of the total income Y has $r = 7400$ different values.

7.1 *Aggregate characteristics in three italian macro-regions*

Table 2 reports for each geographic area: the arithmetic means $M_\ell(X_j)$ and $M_\ell(Y)$; the shares $\gamma_{.\ell}(X_j) = M_\ell(X_j)/M_\ell(Y)$; the shares $\gamma_{..}(X_j) = M(X_j)/M(Y)$; the synthetic index $V_{.\ell}(Y)$; the sum of the weights $n_{.\ell}$; the relative weights $n_{.\ell}/N$; the shares $T_\ell(Y)/T(Y)$. Table 2 shows that the mean value of the South is very far from the means of the other two italian macro-regions. Table 2 shows that for the sources X_2, X_3 and X_4 the shares $\gamma_{.\ell}(X_j)$ of the three macro-regions are "quite" similar and that their differences with the corresponding shares $\gamma_{..}(X_j)$ of the whole country are negligible. Viceversa, the share of X_1 payroll income in Central Italy is "lower" than those in the other two areas. Note that: $\gamma_{..}(X_j) = \sum_{\ell=1}^{k} \gamma_{.\ell}(X_j) \cdot \frac{T_\ell(Y)}{T(Y)}$. Moreover, Table 2 shows that the South has the greatest inequality, while the North has the lowest one, and the inequality of the whole population is a little bit lower than the one of the South. The synthetic inequality index $V(Y) = 0.4808$ means that in the whole population, on average, the lower mean is equal to the (1-0.4808) $\cdot\, 100 \simeq 52\%$ of the overall mean $M(Y)$.

TABLE 2. - *Some aggregate characteristics for geographic areas. Calculations based on Bank of Italy 2014 sample survey data*

	North $\ell = 1$	Center $\ell = 2$	South $\ell = 3$	Italy
$n_{.\ell}$	3866.979	1643.917	2645.103	8156
$n_{.\ell}/N$	0.4741	0.2015	0.3243	1.0000
$V_{.\ell}(Y)$	0.4542	0.4613	0.4909	0.4808=$V(Y)$
Means of Y and X_j, and: shares $T_\ell(Y)/T(Y)$				
X_1	14530.485	11745.379	9448.636	12321.007
X_2	8531.694	9963.766	7190.296	8385.307
X_3	4208.695	3308.157	2623.172	3512.976
X_4	7130.353	7618.478	4284.773	6305.878
Y	34401.23	32635.78	23546.88	30525.17
$T_\ell(Y)/T(Y)$	0.5343	0.2155	0.2502	1.0000
Shares: $\gamma_{.\ell}(X_j)$ and $\gamma_{..}(X_j)$				
$\gamma_{.\ell}(X_1)$	0.4224	0.3599	0.4013	0.4036=$\gamma_{..}(X_1)$
$\gamma_{.\ell}(X_2)$	0.2480	0.3053	0.3054	0.2747=$\gamma_{..}(X_2)$
$\gamma_{.\ell}(X_3)$	0.1223	0.1014	0.1114	0.1151=$\gamma_{..}(X_3)$
$\gamma_{.\ell}(X_4)$	0.2073	0.2334	0.1820	0.2066=$\gamma_{..}(X_4)$
	1.0000	1.0000	1.0000	1.0000

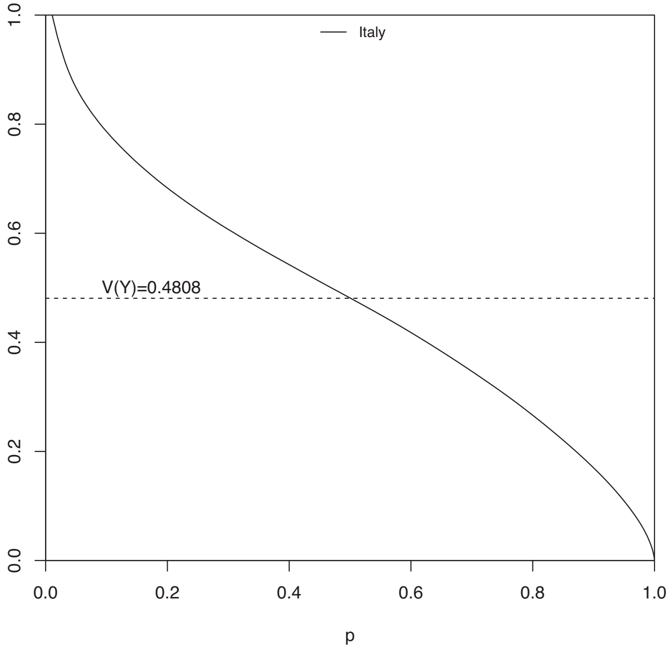
Figure 1 displays two graphs of the point inequality measures for the whole population and the North, the Center and the South. For the subpopulation ℓ the abscissas and the ordinates are given respectively by

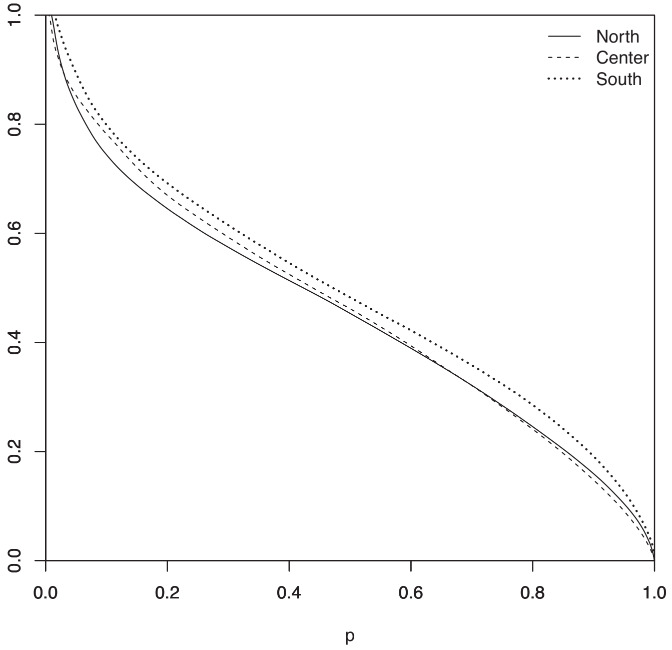
$$p_{h\ell} = \frac{P_{h\ell}}{n_{.\ell}} \quad \text{and} \quad V_{h(p_{h\ell})\ell}(Y) = V_{h\ell}(Y) = \frac{M_\ell(Y) - \bar{M}_{h\ell}(Y)}{M_\ell(Y)},$$

$\forall h = 1, \ldots, r$, while for the whole population the abscissas and the ordinates are given respectively by

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

FIGURE 1. - *Graphs of the point measure for geographic areas. Calculations based on Bank of Italy 2014 sampre survey data*





7.2 *Joint decomposition by geographical areas and sources of the point and synthetic inequality indexes*

The decomposition of $V_h(Y)$ and $V(Y)$ by subpopulations is widely illustrated in Zenga and Valli (2016). In this section we illustrate the decompositions of the point measure$V_{h(p)}(Y)$ for three values of p, and the decompositions of the synthetic index $V(Y) = 0.4808$. For p we have chosen the following values:

- $p = 0.10$, because $V_{h(0.10)}(Y) = 0.7864$ compares the mean income of the poorest 10% households with the mean $M(Y)$;
- $p = 0.50$, because $V_{h(0.50)}(Y) = 0.4812$ compares the mean income of the households with $Y \leq Median(Y)$ with the mean $M(Y)$;
- $p = 0.95$, because $V_{h(0.95)}(Y) = 0.1090$ compares the mean income of the lower group that is the 95% of the whole population with the mean $M(Y)$.

Note that: $h(p) = \left\{\min h : \frac{P_{h.}}{N} \geq p;\ h = 1, \ldots, r\right\}$.

Tables 3, 4 and 5 report the joint decompositions of $V_{h(0.10)}(Y) = 0.7864$, of $V_{h(0.50)}(Y) = 0.4812$ and of $V_{h(0.95)}(Y) = 0.1090$, as well as all the values necessary for their computations. Moreover, Table 6 reports the joint decompositions of the synthetic index $V(Y) = 0.4808$; 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 $V_h(Y)$ with weights $n_{h.}/N$. For example, the value $V_{.3.}(Y) = 0.2498$, reported in Table 6, is the weighted mean of the $r = 7400$ contributions $V_{h3.}(Y) = \frac{M(Y) - \bar{M}_{h3}(Y)}{M(Y)} \cdot p(3|h)$ with weights $n_{h.}/N$. The value of the point index $V_{h(0.10)}(Y) = 0.7864$ means that the mean of the lower group $\bar{M}_{h(0.10).}(Y) = 6521.18$ is equal to the $(1 - 0.7864) \cdot 100 = 21.36\%$ of the mean $M(Y) = 30525.17$. The point index $V_{h(0.50)}(Y) = 0.4812$ means that the mean $\bar{M}_{h(0.50).}(Y) = 15835.87$ is the $(1 - 0.4812) \cdot 100\% = 51.88\%$ of $M(Y) = 30525.17$ and the point index $V_{h(0.95)}(Y) = 0.1090$ informs that the mean of the lower group (that contains the 95% of the whole households) is equal to the $(1 - 0.1090) \cdot 100\% = 89.10\%$ of the income mean of the whole population.

TABLE 3. - *Means and lower means in the geographic areas and joint decomposition of $V_{h(0.1)}(Y) = 0.7864$ by subpopulations and sources. Calculations based on Bank of Italy 2014 sample survey data*

$p = 0.10; h = 440$	North $\ell = 1$	Center $\ell = 2$	South $\ell = 3$	Italy
$Y \leq 10400$	220.68	143.71	451.87	816.26
$Y > 10400$	3646.30	1500.21	2193.235	7339.74
Total $n_{.\ell}$	3866.98	1643.92	2645.10	8156
Conditional relative frequencies				
$p(\ell\|h)$	0.2704	0.1761	0.5536	1.0000
$n_{.\ell}/N$	0.4741	0.2016	0.3243	1.0000
Means and Lower Means				
$\bar{M}_{h\ell}(X_1)$	2540.89	2285.94	1436.80	1884.80
$M_{\ell}(X_1)$	14530.49	11745.38	9448.64	12321.01
$\bar{M}_{h\ell}(X_2)$	2313.97	1939.52	3182.84	2729.04
$M_{\ell}(X_2)$	8531.69	9963.77	7190.30	8385.31
$\bar{M}_{h\ell}(X_3)$	340.83	90.71	361.52	308.25
$M_{\ell}(X_3)$	4208.70	3308.16	2623.17	3512.98
$\bar{M}_{h\ell}(X_4)$	1035.73	2263.58	1662.89	1599.09
$M_{\ell}(X_4)$	7130.35	7618.48	4284.77	6305.88
$\bar{M}_{h\ell}(Y)$	6231.41	6579.75	6644.06	6521.18
$M_{\ell}(Y)$	34401.23	32635.78	23546.88	30525.17
Joint decomposition				
$V_{h(0.1)\ell W}(X_1)$	0.0503	0.0110	0.0474	0.1084=$V_{h(0.1).W}(X_1)$
$V_{h(0.1)\ell B}(X_1)$	0.0363	0.0469	0.1503	0.2335=$V_{h(0.1).B}(X_1)$
$V_{h(0.1)\ell .}(X_1)$	0.0866	0.0579	0.1974	0.3419=$V_{h(0.1)..}(X_1)$
$V_{h(0.1)\ell W}(X_2)$	0.0261	0.0093	0.0236	0.0590=$V_{h(0.1).W}(X_2)$
$V_{h(0.1)\ell B}(X_2)$	0.0277	0.0278	0.0708	0.1263=$V_{h(0.1).B}(X_2)$
$V_{h(0.1)\ell .}(X_2)$	0.0538	0.0372	0.0943	0.1853=$V_{h(0.1)..}(X_2)$
$V_{h(0.1)\ell W}(X_3)$	0.0162	0.0037	0.0133	0.0332=$V_{h(0.1).W}(X_3)$
$V_{h(0.1)\ell B}(X_3)$	0.0119	0.0160	0.0439	0.0718=$V_{h(0.1).B}(X_3)$
$V_{h(0.1)\ell .}(X_3)$	0.0281	0.0197	0.0572	0.1050=$V_{h(0.1)..}(X_3)$
$V_{h(0.1)\ell W}(X_4)$	0.0256	0.0062	0.0154	0.0472=$V_{h(0.1).W}(X_4)$
$V_{h(0.1)\ell B}(X_4)$	0.0211	0.0171	0.0688	0.1070=$V_{h(0.1).B}(X_4)$
$V_{h(0.1)\ell .}(X_4)$	0.0467	0.0233	0.0842	0.1542=$V_{h(0.1)..}(X_4)$
$V_{h(0.1)\ell W}(Y)$	0.1183	0.0303	0.0994	0.2480=$V_{h(0.1).W}(Y)$
$V_{h(0.1)\ell B}(Y)$	0.0969	0.1078	0.3337	0.5384=$V_{h(0.1).B}(Y)$
$V_{h(0.1)\ell .}(Y)$	0.2152	0.1381	0.4331	$V_{h(0.1)}(Y)$=0.7864

TABLE 4. - *Means and lower means in the geographic areas and joint decomposition of $V_{h(0.5)}(Y) = 0.4812$ by subpopulations and sources. Calculations based on Bank of Italy 2014 sample survey data*

$p = 0.50; h = 3262$	North $\ell = 1$	Center $\ell = 2$	South $\ell = 3$	Italy
$Y \leq 25107.88$	1579.84	708.57	1790.70	4079.12
$Y > 25107.88$	2287.14	935.34	854.40	4076.88
Total $n_{.\ell}$	3866.98	1643.92	2645.10	8156
Conditional relative frequencies				
$p(\ell\|h)$	0.3873	0.1737	0.4390	1.0000
$n_{.\ell}/N$	0.4741	0.2016	0.3243	1.0000
Means and Lower Means				
$\bar{M}_{h\ell}(X_1)$	7122.84	5023.37	4812.92	5744.11
$M_\ell(X_1)$	14530.48	11745.38	9448.64.15	12321.01
$\bar{M}_{h\ell}(X_2)$	5629.97	5884.73	6054.45	5860.57
$M_\ell(X_2)$	8531.69	9963.77	7190.30	8385.31
$\bar{M}_{h\ell}(X_3)$	945.77	1089.09	938.66	967.54
$M_\ell(X_3)$	4208.69	3308.16	2623.17	3512.98
$\bar{M}_{h\ell}(X_4)$	3212.64	4163.18	2952.73	3263.66
$M_\ell(X_4)$	7130.35	7618.48	4284.77	6305.88
$\bar{M}_{h\ell}(Y)$	16911.22	16160.37	14758.75	15835.87
$M_\ell(Y)$	34401.23	32635.78	23546.88	30525.17
Joint decomposition				
$V_{h(0.5)\ell W}(X_1)$	0.0077	0.0091	0.0216	0.0739=$V_{h(0.5).W}(X_1)$
$V_{h(0.5)\ell B}(X_1)$	0.0214	0.0338	0.0864	0.1416=$V_{h(0.5).B}(X_1)$
$V_{h(0.5)\ell .}(X_1)$	0.0660	0.0415	0.1080	0.2155=$V_{h(0.5)..}(X_1)$
$V_{h(0.5)\ell W}(X_2)$	0.0175	0.0047	0.0053	0.0275=$V_{h(0.5).W}(X_2)$
$V_{h(0.5)\ell B}(X_2)$	0.0175	0.0096	0.0282	0.0553=$V_{h(0.5).B}(X_2)$
$V_{h(0.5)\ell .}(X_2)$	0.0350	0.0142	0.0335	0.0827=$V_{h(0.5)..}(X_2)$
$V_{h(0.5)\ell W}(X_3)$	0.0196	0.0025	0.0079	0.030=$V_{h(0.5).W}(X_3)$
$V_{h(0.5)\ell B}(X_3)$	0.0129	0.0112	0.0292	0.0533=$V_{h(0.5).B}(X_3)$
$V_{h(0.5)\ell .}(X_3)$	0.0326	0.0138	0.0370	0.0834=$V_{h(0.5)..}(X_3)$
$V_{h(0.5)\ell W}(X_4)$	0.0236	0.0040	0.0062	0.0338=$V_{h(0.5).W}(X_4)$
$V_{h(0.5)\ell B}(X_4)$	0.0157	0.0082	0.0420	0.0659=$V_{h(0.5).B}(X_4)$
$V_{h(0.5)\ell .}(X_4)$	0.0392	0.0122	0.0482	0.0996=$V_{h(0.5)..}(X_4)$
$V_{h(0.5)\ell W}(Y)$	0.1052	0.0189	0.0410	0.1651=$V_{h(0.5).W}(Y)$
$V_{h(0.5)\ell B}(Y)$	0.0675	0.0628	0.1857	0.3161=$V_{h(0.5).B}(Y)$
$V_{h(0.5)\ell .}(Y)$	0.1727	0.0817	0.2267	0.4812=$V_{h(0.5)}(Y)$

TABLE 5. - *Means and lower means in the geographic areas and joint decomposition of $V_{h(0.95)}(Y) = 0.109$ by subpopulations and sources. Calculations based on Bank of Italy 2014 sample survey data*

$p = 0.95$; $h = 6945$	North $\ell = 1$	Center $\ell = 2$	South $\ell = 3$	Italy
$Y \leq 67436.07$	3617.96	1551.71	2579.18	7748.85
$Y > 67436.07$	249.02	92.21	65.92	407.15
Total $n_{.l}$	3866.98	1643.92	2645.10	8156
Conditional relative frequencies				
$p(\ell \mid h)$	0.4669	0.2002	0.3328	1.0000
$n_{.\ell}/N$	0.4741	0.2016	0.3243	1.0000
Means and Lower Means				
$\bar{M}_{h\ell}(X_1)$	13397.66	10925.15	8867.91	11394.82
$M_\ell(X_1)$	14530.48	11745.38	9448.64	12321.01
$\bar{M}_{h\ell}(X_2)$	8084.53	9235.38	7005.38	7955.79
$M_\ell(X_2)$	8531.69	9963.77	7190.30	8385.31
$\bar{M}_{h\ell}(X_3)$	2702.35	2399.14	1772.75	2332.22
$M_\ell(X_3)$	4208.69	3308.16	2623.17	3512.98
$\bar{M}_{h\ell}(X_4)$	6044.88	6859.21	3966.25	5516.08
$M_\ell(X_4)$	7130.35	7618.48	4284.77	6305.88
$\bar{M}_{h\ell}(Y)$	30229.41	29418.88	21612.28	27198.91
$M_\ell(Y)$	34401.23	32635.78	23546.88	30525.17
Joint decomposition				
$V_{h(0.95)\ell W}(X_1)$	0.0082	0.0011	0.0021	$0.0114 = V_{h(0.95).W}(X_1)$
$V_{h(0.95)\ell B}(X_1)$	-0.0247	0.0081	0.0356	$0.0190 = V_{h(0.95).B}(X_1)$
$V_{h(0.95)\ell .}(X_1)$	-0.0165	0.0092	0.0377	$0.0304 = V_{h(0.95)..}(X_1)$
$V_{h(0.95)\ell W}(X_2)$	0.0032	0.0010	0.0007	$0.0049 = V_{h(0.95).W}(X_2)$
$V_{h(0.95)\ell B}(X_2)$	0.0014	-0.0065	0.0144	$0.0093 = V_{h(0.95).B}(X_2)$
$V_{h(0.95)\ell .}(X_2)$	0.0046	-0.0056	0.0150	$0.0140 = V_{h(0.95)..}(X_2)$
$V_{h(0.95)\ell W}(X_3)$	0.0109	0.0012	0.0030	$0.0151 = V_{h(0.95).W}(X_3)$
$V_{h(0.95)\ell B}(X_3)$	0.0015	0.0061	0.0160	$0.0236 = V_{h(0.95).B}(X_3)$
$V_{h(0.95)\ell .}(X_3)$	0.0124	0.0073	0.0190	$0.0387 = V_{h(0.95)..}(X_3)$
$V_{h(0.95)\ell W}(X_4)$	0.0079	0.0010	0.0011	$0.010 = V_{h(0.95).W}(X_4)$
$V_{h(0.95)\ell B}(X_4)$	-0.0039	-0.0046	0.0244	$0.0159 = V_{h(0.95).B}(X_4)$
$V_{h(0.95)\ell .}(X_4)$	0.0040	-0.0036	0.0255	$0.0259 = V_{h(0.95)..}(X_4)$
$V_{h(0.95)\ell W}(Y)$	0.0303	0.0043	0.0068	$0.0413 = V_{h(0.95).W}(Y)$
$V_{h(0.95)\ell B}(Y)$	-0.0257	0.0030	0.0903	$0.0676 = V_{h(0.95).B}(Y)$
$V_{h(0.95)\ell .}(Y)$	0.0045	0.0073	0.0972	$0.1090 = V_{h(0.95)}(Y)$

TABLE 6. - *Joint decomposition of $V(Y) = V_{...}(Y) = 0.4808$. Contributions of within and between components for three regions and four income sources. Calculations based on Bank of Italy 2014 sample survey data*

	North $\ell = 1$	Center $\ell = 2$	South $\ell = 3$	Italy
$V_{.\ell W}(X_1)$	0.0401	0.0071	0.0238	0.0710=$V_{..W}(X_1)$
$V_{.\ell B}(X_1)$	0.0170	0.0315	0.0913	0.1398=$V_{..B}(X_1)$
$V_{.\ell.}(X_1)$	0.0571	0.0386	0.1152	0.2109=$V_{...}(X_1)$
$V_{.\ell W}(X_2)$	0.0165	0.0048	0.0089	0.0302=$V_{..W}(X_2)$
$V_{.\ell B}(X_2)$	0.0164	0.0098	0.0358	0.0620=$V_{..B}(X_2)$
$V_{.\ell.}(X_2)$	0.0329	0.0146	0.0447	0.0922=$V_{...}(X_2)$
$V_{.\ell W}(X_3)$	0.0172	0.0026	0.0083	0.0281=$V_{..W}(X_3)$
$V_{.\ell B}(X_3)$	0.0184	0.0113	0.0302	0.0519=$V_{..B}(X_3)$
$V_{.\ell.}(X_3)$	0.0276	0.0139	0.0385	0.080=$B_{...}(X_3)$
$V_{.\ell W}(X_4)$	0.0213	0.0039	0.0072	0.0324=$V_{..W}(X_4)$
$V_{.\ell B}(X_4)$	0.0134	0.0078	0.0443	0.0655=$V_{..B}(X_4)$
$V_{.\ell.}(X_4)$	0.0347	0.0116	0.0514	0.0977=$V_{...}(X_4)$
$V_{.\ell W}(Y)$	0.0950	0.0183	0.0482	0.1615=$V_{..W}(Y)$
$V_{.\ell B}(Y)$	0.0572	0.0604	0.2016	0.3192=$V_{..B}(Y)$
$V_{.\ell.}(Y)$	0.1523	0.0787	0.2498	0.4808=$V(Y)$

7.2.1 *Macroregions contributions to the point $V_h(Y)$ and synthetic $V(Y)$ indexes*

Now we illustrate the decomposition of the point index $V_{h(0.10)}(Y) = 0.7864$ into the three contribution $V_{h(0.1)\ell.}(Y)$ of each macro region:

$$V_{h(0.1)1.}(Y) = \left[\frac{M(Y) - \bar{M}_{h1}(Y)}{M(Y)}\right] \cdot p(1|h) = 0.7958 \cdot 0.2704 = 0.2152;$$

$$V_{h(0.1)2.}(Y) = \left[\frac{M(Y) - \bar{M}_{h2}(Y)}{M(Y)}\right] \cdot p(2|h) = 0.7844 \cdot 0.1761 = 0.1381;$$

$$V_{h(0.1)3.}(Y) = \left[\frac{M(Y) - \bar{M}_{h3}(Y)}{M(Y)}\right] \cdot p(3|h) = 0.7823 \cdot 0.5536 = 0.4331.$$

These values show that the relative variations of the lower means of the three macro-regions w.r.t the mean of the whole population are similar and very close to the value (0.7864) of the corresponding point index $V_{h(0.1)}(Y)$, while their relative weights $p(\ell|h)$ are very different. This explains why there are so remarkable differences among these three contributions. In particular we note that ''the number'' of the households of the South with $Y \leq y_{h(0.10)} = 10400$ Euro is the 55.36% of the ''number'' of the corresponding households of the whole lower group. This explains why the greatest contribution to the point index $V_{h(0.10)}(Y) = 0.7864$ comes from the South.

The three contributions to the point index $V_{h(0.5)}(Y)$ are:

$$V_{h(0.5)1.}(Y) = \left[\frac{M(Y) - \bar{M}_{h1}(Y)}{M(Y)}\right] \cdot p(1|h) = 0.4460 \cdot 0.3873 = 0.1727;$$

$$V_{h(0.5)2.}(Y) = \left[\frac{M(Y) - \bar{M}_{h2}(Y)}{M(Y)}\right] \cdot p(2|h) = 0.4706 \cdot 0.1737 = 0.0817;$$

$$V_{h(0.5)3.}(Y) = \left[\frac{M(Y) - \bar{M}_{h3}(Y)}{M(Y)}\right] \cdot p(3|h) = 0.5165 \cdot 0.4390 = 0.2267.$$

These values show that the differences between the three relative variations of the lower means $\bar{M}_{h\ell}$ w.r.t. the mean $M(Y)$ and the value (0.4812) of the corresponding point index $V_{h(0.5)}(Y)$ are not remarkable. The relative weights $p(3|h)$ of the South is still greater than the one of the North. Thus, the greatest contribution to $V_{h(0.5)}(Y) = 0.4812$ comes from the South.

Concerning the contribution of the subpopulation ℓ to the point inequality index $V_{h(0.95)}(Y) = 0.1090$ we have:

$$V_{h(0.95)1.}(Y) = \left[\frac{M(Y) - \bar{M}_{h1}(Y)}{M(Y)}\right] \cdot p(1|h) = 0.0097 \cdot 0.4669 = 0.0045;$$

$$V_{h(0.95)2.}(Y) = \left[\frac{M(Y) - \bar{M}_{h2}(Y)}{M(Y)}\right] \cdot p(2|h) = 0.0362 \cdot 0.2002 = 0.0073;$$

$$V_{h(0.95)3.}(Y) = \left[\frac{M(Y) - \bar{M}_{h3}(Y)}{M(Y)}\right] \cdot p(3|h) = 0.2920 \cdot 0.3328 = 0.0972.$$

For $p = 0.95$ the weights $\{p(\ell|h(0.95)) : 0.4669, 0.2002, 0.3328\}$ are very close to the corresponding relative frequencies $\{n_{.\ell}/N : 0.4741, 0.2016, 0.3243\}$, while the relative variations of the lower means $\bar{M}_{h\ell}(Y)$ of the three macro-regions w.r.t. the mean $M(Y)$ of the whole population are very different. In particular, the relative variation of the South is much greater than those of the other two macro-regions. This explains why the greatest contribution of the point index $V_{h(0.95)}(Y)$ comes from the South. In conclusion, in the present application, $V_{h(p)1.}(Y) < V_{h(p)3.}(Y)$ for $p = \{0.1,\ 0.5\ , 0.95\}$. The contribution of the subpopulation ℓ to the synthetic index $V(Y)$ is given by $V_{.\ell.}(Y) = \sum_{h=1}^{r} V_{h\ell.}(Y) \cdot \frac{n_{h.}}{N}$. This explains why the greatest contribution to the synthetic index $V(Y) = 0.4808$ comes from the South (see Table 6).

Table 7 reports for the three macro-regions their:

- relative contributions to the point indexes

$$\nu_{h\ell.}(Y) = \frac{V_{h\ell.}(Y)}{V_h(Y)} = \frac{M(Y) - \bar{M}_{h\ell}(Y)}{M(Y) - \bar{M}_{h.}(Y)} \cdot p(\ell|h), \tag{71}$$

- relative contributions to the synthetic index

$$\nu_{.\ell.}(Y) = \frac{V_{.\ell.}(Y)}{V(Y)} = \sum_{h=1}^{r} \nu_{h\ell.}(Y) \cdot \frac{V_h(Y) \cdot n_{h.}}{V(Y) \cdot N} \tag{72}$$

- relative weights $n_{.\ell}/N$.

TABLE 7. - *Subpopulations relative contributions to the point indexes* $\nu_{h\ell.}(Y) = V_{h\ell.}(Y)/V_h(Y)$ *and to the synthetic index* $\nu_{.\ell.}(Y) = V_{.\ell.}(Y)/V(Y)$. *Calculations based on Bank of Italy 2014 sample survey*

	North	Center	South	
	$\ell = 1$	$\ell = 2$	$\ell = 3$	
$p = 0.10; h = 440$				
$\frac{M(Y) - \bar{M}_{h\ell}(Y)}{M(Y) - \bar{M}_{h.}(Y)}$	1.0121	0.9976	0.9949	
$p(\ell \mid h)$	0.2704	0.1761	0.5536	
$\nu_{h\ell.}(Y) = \frac{M(Y) - \bar{M}_{h\ell}(Y)}{M(Y) - \bar{M}_{h.}(Y)} \cdot p(\ell \mid h)$	0.2736	0.1756	0.5507	1.00
$p = 0.50; h = 3262$				
$\frac{M(Y) - \bar{M}_{h\ell}(Y)}{M(Y) - \bar{M}_{h.}(Y)}$	0.9268	0.9779	1.0733	
$p(\ell \mid h)$	0.3873	0.1737	0.4390	
$\nu_{h\ell.}(Y) = \frac{M(Y) - \bar{M}_{h\ell}(Y)}{M(Y) - \bar{M}_{h.}(Y)} \cdot p(\ell \mid h)$	0.3589	0.1699	0.4712	1.00
$p = 0.95; h = 6945$				
$\frac{M(Y) - \bar{M}_{h\ell}(Y)}{M(Y) - \bar{M}_{h.}(Y)}$	0.0889	0.3326	2.6795	
$p(\ell \mid h)$	0.4669	0.2002	0.3328	
$\nu_{h\ell.}(Y) = \frac{M(Y) - \bar{M}_{h\ell}(Y)}{M(Y) - \bar{M}_{h.}(Y)} \cdot p(\ell \mid h)$	0.0415	0.0666	0.8919	1.00
Synthetic index				
$\nu_{.\ell.}(Y) = \sum_{h=1}^{r} \nu_{h\ell.}(Y) \cdot \frac{V_h(Y) \cdot n_{h.}}{V(Y) \cdot N}$	0.3167	0.1638	0.5195	1.00
$n_{.\ell}/N$	0.4741	0.2016	0.3243	1.00

7.2.2 *Within part of the subpopulations contributions and of the point and synthetic indexes*

The last three rows of Tables 3, 4, 5 and 6 report for the point indexes $V_h(Y)$ and the synthetic index $V(Y)$ the joint decompositions by subpopulations and the within and between parts.

Let us consider now, in detail, the within component: of each contribution $V_{h(p)\ell.}(Y)$, and of the three point indexes $V_{h(p)}(Y)$, for $p = 0.10, 0.50, 0.95$. The within component $V_{h(p)\ell W}(Y)$ of subpopulation ℓ is equal to the relative difference between its mean and lower mean $\frac{M_\ell(Y)-\bar{M}_{h\ell}(Y)}{M(Y)}$ multiplied by the relative frequencies $p(\ell|h)$ and $n_{.\ell}/N$. For each macroregion Table 8 reports the calculations for the three within contributions $V_{h(p)\ell W}(Y)$, and the values of the within contributions $V_{.\ell W}(Y) = \sum_{h=1}^{r} V_{h\ell W}(Y) \cdot \frac{n_{h.}}{N}$. Table 8 shows that the values of the weights $p(\ell|h) \cdot \frac{n_{.\ell}}{N}$ have a considerable influence on the values of the within components: this explains why the North has the greatest within component. Note that the within component of the point index $V_h(Y)$ is given by:

$$V_{h.W}(Y) = \sum_{\ell=1}^{3} V_{h\ell W}(Y) = \sum_{\ell=1}^{3} \frac{M_\ell(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.\ell}}{N}.$$

The percentages (shares) of the contribution $V_{h\ell.}(Y)$ and of the point index $V_h(Y)$, due to the corresponding within components $V_{h\ell W}(Y)$ and $V_{h.W}(Y)$, are respectively given by:

$$\nu_{h\ell W}(Y) = V_{h\ell W}(Y)/V_{h\ell.}(Y) = \frac{M_\ell(Y) - \bar{M}_{h\ell}(Y)}{M(Y) - \bar{M}_{h\ell}(Y)} \cdot \frac{n_{.\ell}}{N}, \tag{73}$$

and

$$\nu_{h.W}(Y) = V_{h.W}(Y)/V_h(Y) = \sum_{\ell=1}^{k} \nu_{h\ell W}(Y) \cdot \nu_{h\ell.}(Y). \tag{74}$$

Obviously, the last share $\nu_{h.W}(Y)$ is also given by:

$$\nu_{h.W}(Y) = \frac{\sum_{\ell=1}^{k} [M_\ell(Y) - \bar{M}_{h\ell}(Y)] \cdot p(\ell|h) \cdot \frac{n_{.\ell}}{N}}{\sum_{\ell=1}^{k} \sum_{g=1}^{k} \left[M_g(Y) - \bar{M}_{h\ell}(Y)\right] \cdot p(\ell|h) \cdot \frac{n_{.g}}{N}}. \tag{75}$$

Finally, the shares of the contribution $V_{.\ell.}(Y)$ and of the synthetic index $V(Y)$, due to the corresponding within components $V_{.\ell W}(Y)$ and $V_{..W}(Y)$, are respectively given by:

$$\nu_{.\ell W}(Y) = V_{.\ell W}(Y)/V_{.\ell.}(Y) = \sum_{h=1}^{r} \nu_{h\ell W}(Y) \cdot \frac{V_h(Y) \cdot n_{h.}}{V(Y) \cdot N} \tag{76}$$

and

$$\nu_{..W}(Y) = V_{..W}(Y)/V(Y) = \sum_{\ell=1}^{k} \nu_{.\ell W}(Y) \cdot \nu_{.\ell.}(Y). \tag{77}$$

By the use of (24) and (29) the last share $\nu_{..W}(Y) = V_{..W}(Y)/V(Y)$ is also given by:

$$\nu_{..W}(Y) = \frac{\sum_{\ell=1}^{k} \sum_{h=1}^{r} \left[\frac{M_\ell(Y)-\bar{M}_{h\ell}(Y)}{M(Y)}\right] p(\ell|h) \cdot \frac{n_{.\ell}}{N} \cdot \frac{n_{h.}}{N}}{\sum_{\ell=1}^{k} \sum_{g=1}^{k} \sum_{h=1}^{r} \left[\frac{M_g(Y)-\bar{M}_{h\ell}(Y)}{M(Y)}\right] \cdot p(\ell|h) \cdot \frac{n_{.g}}{N} \cdot \frac{n_{h.}}{N}}. \tag{78}$$

TABLE 8. - *Within part contributions $V_{h(p)\ell W}(Y)$ and $V_{.\ell W}(Y)$. Calculations based on Bank of Italy 2014 sample survey data*

		$p(\ell\|h)$	$\frac{n_{\ell}}{N}$	$p(\ell\|h)\cdot\frac{n_{\ell}}{N}$	$\frac{M_{\ell}(Y)-\bar{M}_{h\ell}(Y)}{M(Y)}$	$V_{h(p)\ell W}(Y)$
North	$p=0.10$	0.2784	0.4741	0.1282	0.9228	0.1183
	$p=0.50$	0.3873	0.4741	0.1836	0.5730	0.1052
	$p=0.95$	0.4669	0.4741	0.2214	0.1367	0.0303
					$V_{.1W}(Y)=$	0.0950
Center	$p=0.10$	0.1761	0.2016	0.0355	0.8536	0.0303
	$p=0.50$	0.1737	0.2016	0.0350	0.5397	0.0189
	$p=0.95$	0.2002	0.2016	0.0404	0.1054	0.0043
					$V_{.2W}(Y)=$	0.0183
South	$p=0.10$	0.5536	0.3243	0.1795	0.5537	0.0994
	$p=0.50$	0.4390	0.3243	0.1424	0.2879	0.0410
	$p=0.95$	0.3328	0.3243	0.1079	0.0634	0.0068
					$V_{.3W}(Y)=$	0.0482

Table 9 reports the shares: $\nu_{h\ell W}(Y)$, $\nu_{h.W}(Y)$, $\nu_{.\ell W}(Y)$ and $\nu_{...W}(Y)$. For the North and the Center the shares $\nu_{h\ell W}(Y)$ increase for increasing values of p and their mean values are $\nu_{.\ell W}(Y)=0.6238$ and 0.2325 respectively. Viceversa, for the South the share $\nu_{h3W}(Y)$ decreases for increasing values of p and $\nu_{.3W}(Y)=0.1930$.

For the whole country the within component is the 33.59 % of the synthetic index $V(Y)=0.4808$.

7.2.3 *Joint decomposition by three macro-regions and four sources of $V_h(Y)$ and $V(Y)$*

Tables 3, 4 and 5 report the joint decompositions by four sources (X_1, X_2, X_3, X_4) and by three macro-regions (North, Center and South) of the point indexes: $V_{h(0.10)}(Y)=0.7864$, $V_{h(0.50)}(Y)=0.4812$ and $V_{h(0.95)}(Y)=0.1090$. Moreover, each of these latter Tables reports the decompositions by sources of $V_{h(p)}(Y)$, $(p=0.10, 0.50, 0.95)$ as well as the decompositions by sources of the three contribution $V_{h(p)\ell.}(Y)$, $(\ell=1,2,3)$. Finally, Table 6 reports the joint decomposition of the synthetic index $V(Y)=0.4808$ as well as the decompositions by sources of $V(Y)$ and of the three contributions $V_{.\ell.}(Y)$, $(\ell=1,2,3)$. The relative contribution of X_j to:

- the point index $V_{h(p)}(Y)$ is given by:

$$\nu_{h(p)..}(X_j)=\frac{V_{h(p)..}(X_j)}{V_{h(p)}(Y)}=\frac{M(X_j)-\bar{M}_{h.}(X_j)}{M(Y)-\bar{M}_{h.}(Y)};$$

TABLE 9. - *Within part shares:* $\nu_{h\ell W}(Y) = V_{h\ell W}(Y)/V_{h\ell.}(Y)$ *of* $V_{h\ell.}(Y), \nu_{h.W}(Y) = V_{h.W}/V_h(Y)$ *of* $V_h(Y), \nu_{.\ell W}(Y) = V_{.\ell W}(Y)/V_{.\ell.}(Y)$ *of* $V_{.\ell.}(Y)$ *and* $\nu_{..W}(Y) = V_{..W}/V(Y)$ *of* $V(Y)$*. Calculations based on Bank of Italy 2014 sample survey data*

	North	Center	South	Italy
	1	2	3	$\nu_{h.W}(Y)$
$p = 0.10; h = 440$				
$\frac{M_\ell(Y)-\bar{M}_{h\ell}(Y)}{M(Y)-\bar{M}_{h\ell}(Y)}$	1.1595	1.0881	0.7078	
$n_{.\ell}/N$	0.4741	0.2016	0.3243	
$\nu_{h\ell W}(Y) = \frac{M_\ell(Y)-\bar{M}_{h\ell}(Y)}{M(Y)-\bar{M}_{h\ell}(Y)} \cdot n_{.\ell}/N$	0.5497	0.2193	0.2295	0.3153
$p = 0.50; h = 3262$				
$\frac{M_\ell(Y)-\bar{M}_{h\ell}(Y)}{M(Y)-\bar{M}_{h\ell}(Y)}$	1.2847	1.1469	0.5574	
$n_{.\ell}/N$	0.4741	0.2016	0.3243	
$\nu_{h\ell W}(Y) = \frac{M_\ell(Y)-\bar{M}_{h\ell}(Y)}{M(Y)-\bar{M}_{h\ell}(Y)} \cdot n_{.\ell}/N$	0.6090	0.2312	0.1807	0.3431
$p = 0.95; h = 6945$				
$\frac{M_\ell(Y)-\bar{M}_{h\ell}(Y)}{M(Y)-\bar{M}_{h\ell}(Y)}$	14.1056	2.9078	0.2171	
$n_{.\ell}/N$	0.4741	0.2016	0.3243	
$\nu_{h\ell W}(Y) = \frac{M_\ell(Y)-\bar{M}_{h\ell}(Y)}{M(Y)-\bar{M}_{h\ell}(Y)} \cdot n_{.\ell}/N$	6.6874	0.5862	0.070	0.3789
Synthetic index				
$\nu_{.\ell W}(Y) = \sum_{h=1}^{r} \nu_{h\ell W}(Y) \cdot \frac{V_h(Y) \cdot n_{h.}}{V(Y) \cdot N}$	0.6238	0.2325	0.1930	0.3359=$\nu_{..W}(Y)$

- the synthetic index $V(Y)$ is given by:

$$\nu_{...}(X_j) = \frac{V_{...}(X_j)}{V(Y)};$$

- the contribution $V_{h(p)\ell.}(Y)$ is given by

$$\nu_{h(p)\ell.}(X_j) = \frac{V_{h(p)\ell.}(X_j)}{V_{h(p)\ell.}(Y)};$$

- the contribution $V_{.\ell.}(Y)$ is given by:

$$\nu_{.\ell.}(X_j) = \frac{V_{.\ell.}(X_j)}{V_{.\ell.}(Y)}.$$

The relative contributions $\nu_{h(p)\ell.}(X_j)$ and $\nu_{h(p)..}(X_j)$ are reported in Tables 10, 11 and 12, while the relative contributions $\nu_{.\ell.}(X_j)$ and $\nu_{...}(X_j)$ are reported in Table 13.

Table 10 informs that the relative contribution of the variate X_1 ''pay-roll income'' to the point index $V_{h(0.1)}(Y)$ is $\nu_{h(0.1)..}(X_1) = 0.4348$: this means that the difference between the mean $M(X_1)$ and the lower mean $\bar{M}_{h(0.1).}(X_1)$ is the 43.48% of the difference between the corresponding mean and lower mean of the total income Y. Table 13 shows that the value of the relative contribution of the variate X_1 to the synthetic index $V(Y)$ is $\nu_{...}(X_1) = 0.4386$: this means that, on average, the difference between the mean and the lower mean of X_1 is the 43.86% of the corresponding difference between the mean and the lower mean of Y.

In the case of the joint decomposition of $V_h(Y)$ and of $V(Y)$ it is worth:

– to check if the shares $\nu_{h(p)\ell.}(X_j) = V_{h(p)\ell.}(X_j)/V_{h(p)\ell.}(Y)$ are influenced by the subpopulations, and by the cumulative frequency p;

– to find some relationships between the share $\nu_{h(p)..}(X_j)$ and the k shares $\nu_{h(p)\ell.}(X_j)$, as well as between the share $\nu_{...}(X_j)$ and the k shares $\nu_{.\ell.}(X_j)$. From (23) and (51) we obtain:

$$\nu_{h\ell.}(X_j) = \frac{V_{h\ell.}(X_j)}{V_{h\ell.}(Y)} = \frac{M(X_j) - \bar{M}_{h\ell}(X_j)}{M(Y) - \bar{M}_{h\ell}(Y)}. \tag{79}$$

Now, the use of relations $V_{h..}(X_j) = \sum_{\ell=1}^{k} V_{h\ell.}(X_j)$ and $V_h(Y) = \sum_{\ell=1}^{k} V_{h\ell.}(Y)$ in (63) gives:

$$\nu_{h..}(X_j) = \frac{V_{h..}(X_j)}{V_h(Y)} = \frac{\sum_{\ell=1}^{k} V_{h\ell.}(X_j)}{\sum_{\ell=1}^{k} V_{h\ell.}(Y)} = \frac{1}{V_h(Y)} \cdot \sum_{\ell=1}^{k} \frac{V_{h\ell.}(X_j)}{V_{h\ell.}(Y)} V_{h\ell.}(Y).$$

Thus,

$$\nu_{h..}(X_j) = \sum_{\ell=1}^{k} \frac{V_{h\ell.}(X_j)}{V_{h\ell.}(Y)} \cdot \frac{V_{h\ell.}(Y)}{V_h(Y)} = \sum_{\ell=1}^{k} \nu_{h\ell.}(X_j) \cdot \nu_{h\ell.}(Y). \tag{80}$$

In other words, the share $\nu_{h..}(X_j)$ is the ''weighted'' mean of the k shares $\nu_{h\ell.}(X_j)$ with ''weights'' $\nu_{h\ell.}(Y) = \frac{M(Y) - \bar{M}_{h\ell}(Y)}{M(Y) - \bar{M}_{h.}(Y)} \cdot p(\ell|h)$.

For the synthetic index the analogous of (80) is:

$$\nu_{...}(X_j) = \sum_{\ell=1}^{k} \frac{V_{.\ell.}(X_j)}{V_{.\ell.}(Y)} \cdot \frac{V_{.\ell.}(Y)}{V(Y)} = \sum_{\ell=1}^{k} \nu_{.\ell.}(X_j) \cdot \nu_{.\ell.}(Y). \tag{81}$$

Note that the values of the shares $\nu_{h(p)..}(X_j)$ and $\nu_{h(p)\ell.}(X_j)$ reported in Tables 10, 11 and 12 are ''consistent'' with relation (80), the values of the shares $\nu_{...}(X_j)$ and $\nu_{.\ell.}(X_j)$ reported in Table13 are coherent with relation (81).

Besides that, for the variate X_2 (pensions and net transfers) and X_3 (self employment income) the shares $V_{.\ell.}(X_2)$ and $V_{.\ell.}(X_3)$ of the three macro-regions are similar and consequently their differences with the corresponding relative contributions $\nu_{...}(X_2)$ and $\nu_{...}(X_3)$ of the whole country are negligible. We end this section observing that the values of the relative contributions $\nu_{...}(X_j)$ evaluated in this paper (data supplied by the 2014 Central Bank of Italy sample survey) are very similar to the ones obtained in Zenga (2013), Table 12, regarding the 2008 Central Bank of Italy survey.

TABLE 10. - *Shares $\nu_{h(0.1)\ell.}(X_j)$ of income sources to the region contributions $V_{h(0.1)\ell.}(Y)$ and shares $\nu_{h(0.1)..}(X_j)$ of income sources to the point index $V_{h(0.1)}(Y)$. Calculations based on Bank of Italy 2014 sample survey data*

$p = 0.10$; $h = 440$	North	Center	South	Italy
	$\ell = 1$	$\ell = 2$	$\ell = 3$	
$\nu_{h(0.1)\ell.}(X_1)$	0.4026	0.4191	0.4558	0.4348=$\nu_{h(0.1)..}(X_1)$
$\nu_{h(0.1)\ell.}(X_2)$	0.2499	0.2692	0.2178	0.2356=$\nu_{h(0.1)..}(X_2)$
$\nu_{h(0.1)\ell.}(X_3)$	0.1306	0.1429	0.1320	0.1335=$\nu_{h(0.1)..}(X_3)$
$\nu_{h(0.1)\ell.}(X_4)$	0.2169	0.1688	0.1944	0.1961=$\nu_{h(0.1)..}(X_4)$
	1.0000	1.0000	1.000	1.0000

TABLE 11. - *Shares $\nu_{h(0.5)\ell.}(X_j)$ of income sources to the region contributions $V_{h(0.5)\ell.}(Y)$ and shares $\nu_{h(0.5)..}(X_j)$ of income sources to the point index $V_{h(0.5)}(Y)$. Calculations based on Bank of Italy 2014 sample survey data*

$p = 0.50$; $h = 3262$	North	Center	South	Italy
	$\ell = 1$	$\ell = 2$	$\ell = 3$	
$\nu_{h(0.5)\ell.}(X_1)$	0.3818	0.5080	0.4762	0.4477= $\nu_{h(0.5)..}(X_1)$
$\nu_{h(0.5)\ell.}(X_2)$	0.2024	0.1741	0.1478	0.1719=$\nu_{h(0.5)..}(X_2)$
$\nu_{h(0.5)\ell.}(X_3)$	0.1886	0.1687	0.1633	0.1733=$\nu_{h(0.5)..}(X_3)$
$\nu_{h(0.5)\ell.}(X_4)$	0.2272	0.1492	0.2127	0.2071=$\nu_{h(0.5)..}(X_4)$
	1.0000	1.0000	1.000	1.0000

TABLE 12. - *Shares $\nu_{h(0.95)\ell.}(X_j)$ of income sources to the region contributions $V_{h(0.95)\ell.}(Y)$ and shares $\nu_{h(0.95)..}(X_j)$ of income sources to the point index $V_{h(0.95)}(Y)$. Calculations based on Bank of Italy 2014 sample survey data*

$p = 0.95$; $h = 6945$	*North*	Center	South	Italy
	$\ell = 1$	$\ell = 2$	$\ell = 3$	
$\nu_{h(0.95)\ell.}(X_1)$.	-3.6403	1.2618	0.3874	0.2784=$\nu_{h(0.95)..}(X_1)$
$\nu_{h(0.95)\ell.}(X_2)$	1.0170	-0.7684	0.1548	0.1291=$\nu_{h(0.95)..}(X_2)$
$\nu_{h(0.95)\ell.}(X_3)$	2.7409	1.0068	0.1952	0.3550=$\nu_{h(0.95)..}(X_3)$
$\nu_{h(0.95)\ell.}(X_4)$	0.8825	-0.5002	0.2625	0.2374=$\nu_{h(0.95)..}(X_4)$
	1.0000	1.0000	1.000	1.0000

TABLE 13. - *Relative contributions:* $\nu_{...}(X_j)$; $\nu_{.\ell.}(X_j)$. *Calculations based on Bank of Italy 2014 sample survey data*

	North	Center	South	Italy
	$\ell = 1$	$\ell = 2$	$\ell = 3$	
$\nu_{.\ell.}(X_1)$	0.3753	0.4900	0.4610	0.4386=$\nu_{...}(X_1)$
$\nu_{.\ell.}(X_2)$	0.2160	0.1857	0.1788	0.1917=$\nu_{...}(X_2)$
$\nu_{.\ell.}(X_3)$	0.1810	0.1765	0.1543	0.1664=$\nu_{...}(X_3)$
$\nu_{.\ell.}(X_4)$	0.2276	0.1479	0.2059	0.2033=$\nu_{...}(X_4)$
	1.0000	1.0000	1.000	1.0000

7.3 *Decomposition by sources for each Italian macro-regions*

This section compares the decompositions by sources of each macro-region with the corresponding decompositions of the whole country. Besides that, the decompositions by sources of the point indexes $V_{h\ell}(Y)$ (of the macro-regions) are compared with the corresponding decompositions of the contributions $V_{h(p)\ell.}(Y)$. Table 14 reports for $p = 0.10$, 0.50 and 0.95 the decomposition by sources of the point indexes $V_{h\ell}(Y)$ and $V_h(Y)$, while Table 15 reports the decomposition by sources of the synthetic indexes $V_{.\ell}(Y)$ and $V(Y)$. This latter table shows that the South has the greatest inequality, the North has the lowest one, and the inequality of the whole population is lower than the one of the South and much greater than those of the other two macro-regions. Note that the value of the synthetic Bonferroni index $V(Y)$ is not necessary between the extremes values of the subpopulations inequalities $V_{.\ell}(Y)$: see Table 9 in Zenga and Valli (2016). Moreover, Table 15 shows that:

i) the relative contributions of the North $\nu_{.1}(X_j)$, $j = 1, 2, 3, 4$ are similar to the corresponding relative contributions $\nu_{...}(X_j)$ of the whole country;
ii) the relative contribution of the Center is much greater than the relative contribution of the whole country for X_2 and it is much lower for X_1;
iii) the relative contributions of the South are very similar to the corresponding relative contributions of the whole country for X_2 and X_3, while for X_1 and X_4 there are remarkable differences.

Table 14 shows that there are important differences between the relative contribution $\nu_{h\ell}(X_j)$ and $\nu_{h..}(X_j)$. Now, it is worth to point out that there are also considerable differences between the relative contributions

$$\nu_{h\ell}(X_j) = \frac{V_{h\ell}(X_j)}{V_{h\ell}(Y)} = \frac{\frac{M_\ell(X_j) - \bar{M}_{h\ell}(X_j)}{M_\ell(Y)}}{\frac{M_\ell(Y) - \bar{M}_{h\ell}(Y)}{M_\ell(Y)}} = \frac{M_\ell(X_j) - \bar{M}_{h\ell}(X_j)}{M_\ell(Y) - \bar{M}_{h\ell}(Y)}, \tag{82}$$

reported in Table (14) and the shares

$$\nu_{h\ell.}(X_j) = \frac{V_{h\ell.}(X_j)}{V_{h\ell.}(Y)} = \frac{\frac{M(X_j) - \bar{M}_{h\ell}(X_j)}{M(Y)} \cdot p(\ell|h)}{\frac{M(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot p(\ell|h)} = \frac{M(X_j) - \bar{M}_{h\ell}(X_j)}{M(Y) - \bar{M}_{h\ell}(Y)}$$

reported in Tables 10, 11 and 12. In (82) we compare the lower means $\bar{M}_{h\ell}(X_j)$ and $\bar{M}_{h\ell}(Y)$ with the means $M_\ell(X_j)$ and $M_\ell(Y)$ of the same subpopulation ℓ, while in (79) we compare the lower means $\bar{M}_{h\ell}(X_j)$ and $\bar{M}_{h\ell}(Y)$ of the subpopulation ℓ with the means $M(X_j)$ and $M(Y)$ of the whole population. Obviously, the values of $M(Y) = \sum_{g=1}^{k} M_g(Y) \cdot \frac{n_{.g}}{N}$ and of $M(X_j) = \sum_{g=1}^{k} M_g(X_j) \cdot \frac{n_{.g}}{N}$ are influenced by the values of all the k means $M_g(X_j)$ and $M_g(Y)$ as well as by the corresponding relative frequencies $\frac{n_{.g}}{N}$. This explains why we may have important differences between $\nu_{h\ell}(X_j)$ and $\nu_{h\ell.}(X_j)$.

Finally, we end this section showing that the relative contribution $\nu_{h\ell}(X_j)$ of X_j to the point index $V_{h\ell}(Y)$ of the subpopulation ℓ is equal to the ratio

$$\nu_{h\ell W}(X_j) = \frac{V_{h\ell W}(X_j)}{V_{h\ell W}(Y)}. \tag{83}$$

TABLE 14. - *Decomposition by sources of the point indexes $V_{h\ell}(Y)$ and $V_h(Y)$. Calculations based on Bank of Italy 2014 sample survey data*

	North		Center		South		Italy	
	$V_{h1}(\cdot)$	$\nu_{h1}(\cdot)$	$V_{h2}(\cdot)$	$\nu_{h2}(\cdot)$	$V_{h3}(\cdot)$	$\nu_{h3}(\cdot)$	$V_{h..}(\cdot)$	$\nu_{h..}(\cdot)$
$p = 0.10; h = 440; Y \leq 10400.00$								
X_1	0.3485	*0.4256*	0.2898	*0.3630*	0.3403	*0.4740*	0.3419	*0.4348*
X_2	0.1807	*0.2207*	0.2459	*0.3080*	0.1702	*0.2371*	0.1853	*0.2356*
X_3	0.1124	*0.1373*	0.0986	*0.1235*	0.0960	*0.1338*	0.1050	*0.1335*
X_4	0.1772	*0.2164*	0.1641	*0.2055*	0.1113	*0.1551*	0.1542	*0.1961*
$V_{h\ell}(Y)$	0.8189	*1.0000*	0.7984	*1.0000*	0.7178	*1.0000*	0.7864	*1.0000*
$p = 0.50; h = 3262; Y \leq 25107.88$								
X_1	0.2153	*0.4235*	0.2060	*0.4080*	0.1969	*0.5275*	0.2155	*0.4477*
X_2	0.0843	*0.1659*	0.1250	*0.2476*	0.0482	*0.1292*	0.0827	*0.1719*
X_3	0.0948	*0.1866*	0.0680	*0.1347*	0.0715	*0.1917*	0.0834	*0.1733*
X_4	0.1139	*0.2240*	0.1059	*0.2097*	0.0566	*0.1516*	0.0997	*0.2071*
$V_{h\ell}(Y)$	0.5084	*1.0000*	0.5048	*1.0000*	0.3732	*1.0000*	0.4812	*1.0000*
$p = 0.95; h = 6945; Y \leq 67436.07$								
X_1	0.0329	*0.2715*	0.0251	*0.2550*	0.0247	*0.3002*	0.0303	*0.2784*
X_2	0.0130	*0.1072*	0.0223	*0.2265*	0.0079	*0.0956*	0.0141	*0.1291*
X_3	0.0438	*0.3611*	0.0279	*0.2826*	0.0361	*0.4396*	0.0387	*0.3550*
X_4	0.0316	*0.2602*	0.0233	*0.2360*	0.0135	*0.1646*	0.0259	*0.2374*
$V_{h\ell}(Y)$	0.1213	*1.0000*	0.0986	*1.0000*	0.0821	*1.0000*	0.1090	*1.0000*

In (83) $V_{h\ell W}(X_j)$ and $V_{h\ell W}(Y)$ are defined as follow (see Section 6):

$$V_{h\ell W}(X_j) = V_{h\ell\ell}(X_j) = \frac{M_\ell(X_j) - \bar{M}_{h\ell}(X_j)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.\ell}}{N},$$

$$V_{h\ell W}(Y) = V_{h\ell\ell}(Y) = \frac{M_\ell(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.\ell}}{N}.$$

Note that $V_{h\ell W}(X_j)$ and $V_{h\ell W}(Y)$ are the within parts of $V_{h\ell.}(X_j)$ and $V_{h\ell.}(Y)$ respectively.

In conclusion, by the use of (25) and (53) in (83) we get the result:

$$\nu_{h\ell W}(X_j) = \frac{V_{h\ell W}(X_j)}{V_{h\ell W}(Y)} = \frac{M_\ell(X_j) - \bar{M}_{h\ell}(X_j)}{M_\ell(Y) - \bar{M}_{h\ell}(Y)} = \nu_{h\ell}(X_j). \quad (84)$$

TABLE 15. - *Decomposition by sources of the synthetic indexes $V_{.\ell}(Y)$ and $V(Y)$. Calculations based on Bank of Italy 2014 sample survey data*

	North		Center		South		Italy	
	$V_{.1}(\cdot)$	$\nu_{.1}(\cdot)$	$V_{.2}(\cdot)$	$\nu_{.2}(\cdot)$	$V_{.3}(\cdot)$	$\nu_{.3}(\cdot)$	$V_{...}(\cdot)$	$\nu_{...}(\cdot)$
X_1	0.1918	*0.4224*	0.1789	*0.3877*	0.2426	0.4942	0.2109	*0.4386*
X_2	0.0767	*0.1688*	0.1197	*0.2595*	0.0957	0.1949	0.0922	*0.1917*
X_3	0.0835	*0.1838*	0.0650	*0.1410*	0.0798	0.1625	0.0800	*0.1664*
X_4	0.1022	*0.2250*	0.0977	*0.2177*	0.0728	0.1484	0.0977	*0.2033*
$V_{.\ell}(Y)$	0.4542	*1.0000*	0.4613	*1.0000*	*0.4909*	1.0000	0.4808	*1.0000*

8. PRINCIPAL RESULTS

In the case of frequency distribution framework $\{0 \le y_1 < \ldots < y_h < \ldots < y_r\}$ is the set of the distinct values assumed by the total income Y over the k subpopulations and $\{n_{1.}, \ldots, n_{h.}, \ldots, n_{r.}; \sum_{h=1}^{r} n_{h.} = N\}$ are the corresponding frequencies, N is the size of the whole population. In this case, see Zenga and Valli (2016), the Bonferroni point $V_h(Y)$ and synthetic $V(Y)$ indexes are given by:

$$V_h(Y) = \frac{M(Y) - \bar{M}_{h.}(Y)}{M(Y)} \quad (4) \quad \text{and} \quad V(Y) = \sum_{h=1}^{r} V_h(Y) \cdot \frac{n_{h.}}{N}.$$

In (4) $M(Y)$ is the mean of total income Y and $\bar{M}_{h.}(Y)$ is the lower mean (the mean of Y computed on the $P_{h.} = \sum_{t=1}^{h} n_{t.}$ units of the lower group $\{Y \le y_h\}$).

1) In the decomposition by sources the total income Y is the sum of c sources $(X_1, \ldots, X_j, \ldots, X_c)$ observable on each of the N units of the population. Obviously,

$$M(Y) = \sum_{j=1}^{c} M(X_j) \quad \text{and} \quad \bar{M}_{h.}(Y) = \sum_{j=1}^{c} \bar{M}_{h.}(X_j),$$

where $M(X_j)$ is the mean of X_j in the whole population and $\bar{M}_{h.}(X_j)$ is the mean (the lower mean) of X_j in the lower group. Using (first step) relations (8) in the numerator of $V_h(Y)$ the following additive decomposition by sources of $V_h(Y)$ is obtained:

$$V_h(Y) = \sum_{j=1}^{c} V_h(X_j), \quad \text{where} \quad V_h(X_j) = \frac{M(X_j) - \bar{M}_{h.}(X_j)}{M(Y)}$$

is the contribution of X_j to $V_h(Y)$. Finally, putting (second step) the decomposition (9) in (5), the decomposition by sources of $V(Y)$ is obtained:

$$V(Y) = \sum_{j=1}^{c} V(X_j), \quad \text{where} \quad V(X_j) = \sum_{h=1}^{r} V_h(X_j) \cdot \frac{n_{h.}}{N}$$

is the contribution of X_j to the synthetic index $V(Y)$.

2) In the decomposition by subpopulations we need the bivariate distribution of the N units according to k disjoint subpopulations and the r distinct values of Y. In this distribution (see Table 1): n_{hg} denotes the frequency of the value y_h in the subpopulation g; $n_{h.} = \sum_{g=1}^{k} n_{hg}$; $n_{.g} = \sum_{h=1}^{r} n_{hg}$ is the size of subpopulation g; $P_{hg} = \sum_{t=1}^{h} n_{tg}$ are the units of subpopulation g with $Y \le y_h$; $P_{h.} = \sum_{g=1}^{k} P_{hg}$; $N = \sum_{g=1}^{k} n_{.g} = \sum_{h=1}^{r} n_{h.}$. Moreover: $M_g(Y)$ and $\bar{M}_{hg}(Y)$ are the mean and the lower mean of subpopulation g, $\frac{n_{.g}}{N}$ is the relative frequency of subpopulation g and $p(\ell|h) = \frac{P_{hg}}{P_{h.}}$ is the relative frequency of subpopulation g in the lower group. By the use of Property II (Section 2):

$$M(Y) = \sum_{g=1}^{k} M_g(Y) \cdot \frac{n_{.g}}{N} \quad \text{and} \quad \bar{M}_{h.}(Y) = \sum_{\ell=1}^{k} \bar{M}_{h\ell}(Y) \cdot p(\ell|h).$$

Using (first step) these two latter relations in the numerator of (4), Zenga and Valli (2016) have obtained the following basic $k \times k$ decomposition of $V_h(Y)$:

$$V_h(Y) = \sum_{g=1}^{k} \sum_{\ell=1}^{k} V_{h\ell g}(Y), \quad V_{h\ell g}(Y) = \frac{M_g(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.g}}{N}$$

is the contribution to the point inequality index $V_h(Y)$ that derives from the comparison of the lower mean $\bar{M}_{h\ell}(Y)$ with the mean $M_g(Y)$. Obviously, $\sum_{\ell=1}^{k} \sum_{g=1}^{k} p(\ell|h) \cdot \frac{n_{.g}}{N} = \sum_{\ell=1}^{k} p(\ell|h) \cdot \sum_{g=1}^{k} \frac{n_{.g}}{N} = 1 \cdot 1 = 1$. It is worth to remark that, from the basic $k \times k$ decomposition (21), through different aggregations of the contributions $V_{h\ell g}(Y)$, many others additive decompositions of $V_h(Y)$ are derived. In particular, $V_h(Y)$ is decomposed in the sums (24) and (26).

$$V_h(Y) = \sum_{\ell=1}^{k} V_{h\ell.}(Y)(24), \text{where}$$

$$V_{h\ell.}(Y) = \sum_{g=1}^{k} V_{h\ell g}(Y) = \frac{M(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot p(\ell|h)$$

is the contribution of subpopulation ℓ to $V_h(Y)$.

$$V_h(Y) = V_{h.W}(Y) + V_{h.B}(Y) \ (26), \text{ where}$$

$$V_{h.W}(Y) = \sum_{\ell=1}^{k} V_{h\ell\ell}(Y) = \sum_{\ell=1}^{k} \frac{M_\ell(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.\ell}}{N}$$

and

$$V_{h.B}(Y) = \sum_{\ell=1}^{k} \sum_{g \neq \ell} V_{h\ell g}(Y) = \sum_{\ell=1}^{k} \sum_{g \neq \ell} \frac{M_g(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.g}}{N}$$

are the within and the between components of $V_h(Y)$, respectively.

Putting (second step) the $k \times k$ decomposition (21) of $V_h(Y)$ in (5) the following $k \times k$ additive decomposition of $V(Y)$ is obtained:

$$V(Y) = \sum_{\ell=1}^{k} \sum_{g=1}^{k} V_{.\ell g}(Y) \ (27), \quad \text{where} \quad V_{.\ell g}(Y) = \sum_{h=1}^{r} V_{h\ell g}(Y) \cdot \frac{n_{h.}}{N}.$$

Then, in the usual way, $V(Y)$ is decomposed in the sums (28) and (30).

$$V(Y) = \sum_{\ell=1}^{k} V_{.\ell.}(Y) \ (28), \quad \text{where} \quad V_{.\ell.}(Y) = \sum_{g=1}^{k} V_{.\ell g}(Y)$$

is the contribution of subpopulation ℓ to the synthetic index $V(Y)$.

$$V(Y) = V_{..W}(Y) + V_{..B}(Y) \ (30), \text{ where}$$

$$V_{..W}(Y) = \sum_{\ell=1}^{k} \left[\sum_{h=1}^{r} \frac{M_\ell(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.\ell}}{N} \cdot \frac{n_{h.}}{N} \right]$$

and

$$V_{..B}(Y) = \sum_{\ell=1}^{k} \sum_{g \neq \ell}^{k} \left[\sum_{h=1}^{r} \frac{M_g(Y) - \bar{M}_{h\ell}(Y)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.g}}{N} \cdot \frac{n_{h.}}{N} \right]$$

are the within and the between parts of $V(Y)$, respectively.

3) Using (first step) the relations $\bar{M}_{h\ell}(Y) = \sum_{j=1}^{c} \bar{M}_{h\ell}(X_j)$ and $M_g(Y) = \sum_{j=1}^{c} M_g(X_j)$ in the $k \times k$ decomposition (by subpopulations) (21) of $V_h(Y)$, the present paper obtains the following $k \times k \times c$ joint decomposition, by subpopulations and sources of $V_h(Y)$:

$$V_h(Y) = \sum_{\ell=1}^{k} \sum_{g=1}^{k} \sum_{j=1}^{c} V_{h\ell g}(X_j), \ (47)$$

where

$$V_{h\ell g}(X_j) = \frac{M_g(X_j) - \bar{M}_{h\ell}(X_j)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.g}}{N},$$

and $M_g(X_j)$ and $\bar{M}_{h\ell}(X_j)$ are the means of X_j in the subpopulation g and in the lower group $(Y \leq y_h)$ of subpopulation ℓ.

Putting (second step) the $k \times k \times c$ joint decomposition (47) of $V_h(Y)$ in (5), the following $k \times k \times c$ joint decomposition of $V(Y)$ is obtained:

$$V(Y) = \sum_{\ell=1}^{k} \sum_{g=1}^{k} \sum_{j=1}^{c} V_{.\ell g}(X_j) \ (59),$$

where

$$V_{.\ell g}(X_j) = \sum_{h=1}^{r} \frac{M_g(X_j) - \bar{M}_{h\ell}(X_j)}{M(Y)} \cdot p(\ell|h) \cdot \frac{n_{.g}}{N} \cdot \frac{n_{h.}}{N}$$

is the weighted arithmetic mean of $V_{h\ell g}(X_j)$ with weights $\frac{n_{h.}}{N}$.

The theoretical results of this paper are applied to the 2014 Bank of Italy sample survey on income and wealth. Section 7 illustrates the decompositions: of the point index $V_{h(p)}(Y) = V_h(Y)$ for three values (0.1, 0.5, 0.95) of the cumulative relative frequency p, and of the synthetic index $V(Y)$. Now, it is worth to remark that, the interpretation of the decomposition above reported can be facilitated by the calculation of some relative contributions. In the case of the decomposition by sources, the relative contribution of X_j to $V_h(Y)$ and to $V(Y)$ are:

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

and

$$\nu_{...}(X_j) = \frac{V_{...}(X_j)}{V(Y)} = \sum_{h=1}^{r} \nu_{h..}(X_j) \cdot \frac{V_h(Y) \cdot n_{h.}}{V(Y) \cdot N}$$

Moreover, the relative contributions $\nu_{h..}(X_j)$ and $\nu_{...}(X_j)$ should be compared with the shares $\gamma_{..}(X_j) = \frac{M(X_j)}{M(Y)}$ in order to discern whether a given income source X_j has an exarcebating or a mitigating impact on inequality in the distribution of total income Y. Table 16 reports the relative contributions $\nu_{h..}(X_j)$ and $\nu_{...}(X_j)$ and the shares $\gamma_{..}(X_j)$.

TABLE 16. - *Shares $\gamma_{..}(X_j)$, and relative contributions $\nu_{h(p)..}(X_j)$ and $\nu_{...}(X_j)$. Calculations based on Bank of Italy 2014 sample survey data*

	$\nu_{h(0.10)..}(\cdot)$	$\nu_{h(0.50)..}(\cdot)$	$\nu_{h(0.95)..}(\cdot)$	$\nu_{...}(\cdot)$	$\gamma_{..}(\cdot)q$
X_1	0.4348	0.4477	0.2789	0.4383	0.4036
X_2	0.2356	0.1719	0.1284	0.1917	0.2747
X_3	0.1335	0.1733	0.3550	0.1664	0.1151
X_4	0.1961	0.2071	0.2376	0.2033	0.2066
	1.00	1.00	1.00	1.00	1.00

Table 16 shows that: for the variate X_2, $\nu_{...}(X_2) < \gamma_{..}(X_2)$; for the variates X_3 and X_1, $\nu_{...}(X_3) > \gamma_{..}(X_3)$ and $\nu_{...}(X_1) > \gamma_{..}(X_1)$. Thus, the pension and net transfert X_2 decreases the inequality, while the net self employment X_3 and the payroll income X_1 increase the inequality.

In Zenga (2013) is shown that: $a)$ the relative contributions of X_j to the Zenga (2007), Bonferroni (1930) and Gini (1914) point indexes are equal; $b)$ the relative contributions of X_j to the corresponding synthetic indexes are the weighted arithmetic means of the relative point contributions, with weights $\frac{I_h(Y)}{I(Y)} \cdot \frac{n_{h.}}{N}$, $\frac{V_h(Y)}{V(Y)} \cdot \frac{n_{h.}}{N}$ and $\frac{G_h(Y)}{G(Y)} \cdot \frac{n_{h.}}{N}$, respectively. For more details on the comparisons of the relative contributions $\nu_{h..}(X_j)$ and $\nu_{...}(X_j)$ w.r.t. $\gamma_{..}(X_j)$ see: Zenga *et al.* (2012), Zenga (2013) and Pasquazzi and Zenga (2018). In the case of the decomposition by subpopulations, the point $V_h(Y)$ and the synthetic Bonferroni indexes can be decomposed in the sum of the k contributions $V_{h\ell.}(Y)$ and $V_{.\ell.}(Y)$, respectively.

The relative contributions of the subpopulation ℓ to $V_h(Y)$ and $V(Y)$ are:

$$\nu_{h\ell.}(Y) = \frac{V_{h\ell.}(Y)}{V_h(Y)} = \frac{M(Y) - \bar{M}_{h\ell}(Y)}{M(Y) - \bar{M}_{h.}(Y)} \cdot p(\ell|h),$$

and

$$\nu_{.\ell.}(Y) = \frac{V_{.\ell.}(Y)}{V(Y)} = \sum_{h=1}^{r} \nu_{h\ell.}(Y) \cdot \frac{V_h(Y) \cdot n_{h.}}{V(Y) \cdot N}$$

Unfortunately, the contributions $V_{h\ell g}(Y)$, $V_{h\ell.}(Y)$ and $V_{.\ell.}(Y)$ can be negative. For more details on this point see Zenga and Valli (2016).

It is worthwhile to remark that in the decomposition by subpopulations of the Gini index, we can have negative contributions, too. For more details see Zenga (2016b). The fact that some contributions are negative may cause some difficulties on the interpretation. This cannot happens for the Zenga (2007) inequality index. In fact, by construction, see Zenga (2016), the contributions

$$B_{h\ell g}(Y) = \frac{\overset{+}{M}_{hg}(Y) - \bar{M}_{h\ell}(Y)}{\overset{+}{M}_{h.}(Y)} \cdot a(g|h) \cdot p(\ell|h)$$

to the Zenga (2007) point inequality index

$$I_h(Y) = \frac{\overset{+}{M}_{h.}(Y) - \bar{M}_{h.}(Y)}{\overset{+}{M}_{h.}(Y)}$$

are never negative; $\overset{+}{M}_{h.}(Y)$ is the mean of the upper group $\{Y > \; y_h\}$ of the whole population, $\overset{+}{M}_{hg}(Y)$ is the mean of the upper group of the subpopulation g and $a(g|h)$ is the relative frequency of subpopulation g in the upper group. In the case of the joint decomposition of the point index $V_h(Y)$ are useful the shares

$$\nu_{h\ell.}(X_j) = \frac{V_{h\ell.}(X_j)}{V_{h\ell.}(Y)} = \frac{M(X_j) - \bar{M}_{h\ell}(X_j)}{M(Y) - \bar{M}_{h\ell}(Y)}$$

and their relationship with $\nu_{h..}(X_j)$:

$$\nu_{h..}(X_j) = \sum_{\ell=1}^{k} \nu_{h\ell.}(X_j) \cdot \nu_{h\ell.}(Y).$$

For the joint decomposition of $V(Y)$ are useful the shares:

$$\nu_{.\ell.}(X_j) = \frac{V_{.\ell.}(X_j)}{V_{.\ell.}(Y)}$$

and their relationship with $\nu_{...}(X_j)$:

$$\nu_{...}(X_j) = \frac{V_{...}(X_j)}{V(Y)} = \sum_{\ell=1}^{k} \nu_{.\ell.}(X_j) \cdot \nu_{.\ell.}(Y).$$

Note that the values of $\nu_{h\ell.}(X_j)$ and of $\nu_{h..}(X_j)$, reported in Tables 10, 11 and 12, are ‘‘consistent’’ with relation (80), and the values of the shares $\nu_{.\ell.}(X_j)$ and $\nu_{...}(X_j)$, reported in Table 13, are coherent with relation (81). Besides that, the shares $\nu_{.\ell.}(X_2)$ and of $\nu_{.\ell.}(X_3)$ of the three macro regions are similar and, consequently, their differences with the corresponding relative contributions $\nu_{...}(X_2)$ and $\nu_{...}(X_3)$ of the whole country are negligible. Moreover, this paper compares in Subsection 7.3 : the decompositions by sources of each macro-region with the corresponding decompositions of the whole country; the decomposition by sources of the point indexes $V_{h\ell}(Y)$ (of the macro region) with the corresponding decompositions by sources of the contributions $V_{h\ell.}(Y)$. In particular, Table 15 shows that only the relative contributions of the North $\nu_{.1}(X_j)$, $j = 1, 2, 3, 4$ are similar to the corresponding relative contributions $\nu_{...}(X_j)$ of the whole country. Table 14 shows that there are important differences: between the relative contributions $\nu_{h\ell}(X_j)$ and $\nu_{h..}(X_j)$, and between the relative contributions $\nu_{h\ell}(X_j)$, reported in Table 14, and the shares $\nu_{h\ell.}(X_j)$, reported in Tables 10, 11 and 12. Finally, Subsection 7.3 shows that the relative contribution $\nu_{h\ell}(X_j)$ of X_j to the point index $V_{h\ell}(Y) = \frac{M_\ell(Y) - \bar{M}_{h\ell}(Y)}{M_\ell(Y)}$ of the subpopulation ℓ is equal to the ratio $\nu_{h\ell W}(X_j) = V_{h\ell W}(X_j)/V_{h\ell W}(Y)$, where $V_{h\ell W}(Y)$ and $V_{h\ell W}(X_j)$ are the within parts of $V_{h\ell.}(Y)$ and $V_{h\ell.}(X_j)$, respectively.

REFERENCES

Bank of Italy (2016). Household income and wealth in 2014. *Supplements to the Statistical Bulletin*, Sample Survey, **64**(XXV).

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

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

Pasquazzi L., Zenga M. (2018). Components of Gini, Bonferroni and Zenga inequality indexes for EU income data. *Journal of 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*, **V**(1), 3-27.

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

Zenga M. (2013). Decomposition by sources of the Gini, Bonferroni and Zenga inequality indexes. *Statistica & Applicazioni*, **XI**(2), 131-158.

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

Zenga M. (2016). On the decomposition by subpopulations of the point and the 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. *Statistica & Applicazioni*, **XIV**(1), 3-28.

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