Statistica & Applicazioni - Vol. XI, n. 1, 2013, pp. 63-77

A LONGITUDINAL DECOMPOSITION OF ZENGA'S NEW INEQUALITY INDEX

Mauro Mussini*
Michele Zenga*

SUMMARY

The paper proposes a three-term decomposition of Zenga's new inequality index over time. Given an initial and a final time, the link among inequality trend, re-ranking, and income growth is explained by decomposing the inequality index at the final time into three components: one measuring the effect of re-ranking between individuals, a second term gauging the effect of disproportional growth between individuals' incomes, and a third component measuring the impact of the inequality existing at the initial time. The decomposition allows one to distinguish the determinants of inequality change from the contribution of the inequality at the initial time to the inequality at the final time. We applied the decomposition to Italian household income data collected by the Survey on Household Income and Wealth of the Bank of Italy, waves 2008-2010.

Keywords: *Zenga Index, Decomposition, Inequality Trend, Re-ranking, Income Growth.*

1. INTRODUCTION

In recent years, research on income inequality measurement has paid growing attention to decomposing inequality change over time (Jenkins and Van Kerm, 2006; O'Neill and Van Kerm, 2008; Wagstaff, 2009). In this type of longitudinal decomposition of inequality, the change in inequality between two points in time is split into two components: one measuring the effect of re-ranking between individuals, the other measuring the effect of disproportional growth between individuals' incomes. The re-ranking component always provides a contribution increasing inequality, whereas the contribution of income growth can increase or reduce inequality over time.

In this paper, we develop a longitudinal decomposition for the Zenga inequality index (Zenga, 2007a). The Zenga index is based on the underlying Zenga inequality curve, each point of which measures the relative income-gap between the mean income of the poorest 100p per cent of population and the mean income of the remaining part of population. The index satisfies various desirable properties (anonymity, scale invariance, population replication, and principle of transfers), and provides straightforward graphic interpretation. Several studies contributed to point out metho-

* Dipartimento di Statistica e Metodi Quantitativi - Università Milano Bicocca - via Bicocca degli Arcimboldi, 8 - 20126 Milano (e-mail: 251658240 ✉ mauro.mussini1@unimib.it, michele.zenga@unimib.it).

dological and empirical aspects relevant to the Zenga index.[1] The index was decomposed by subgroup by Radaelli (2008, 2010), and its source decomposition was proposed by Zenga, Radaelli and Zenga (2011). We contribute to make the index decomposition framework more comprehensive by showing that the Zenga index can be decomposed by time.

Given an initial and a final time, we provide a three-term decomposition of the Zenga index at the final time: a component measuring the effect of re-ranking between individuals, a second term gauging the effect of disproportional growth between individuals' incomes, and a third component measuring the impact of the inequality existing at the initial time. This decomposition approach differs from that of the above mentioned longitudinal decomposition, since the former does not decompose the difference in inequality between the two points in time. In fact, it accounts for the inequality at the initial time by isolating its contribution to the inequality at the final time. Therefore, the decomposition allows one to detect the determinants of inequality trend and, simultaneously, the impact of the inequality at the initial time on the inequality at the final time.

The article is organised as follows. Section 2 briefly reviews the Zenga index and introduces notations. Section 3 develops the longitudinal decomposition of the Zenga index. Section 4 shows a comparison between the decomposition approach based on the Zenga index and that based on the generalised Gini index. In Section 5, the decomposition is applied to Italian household income data collected by the Survey on Household Income and Wealth of the Bank of Italy, waves 2008-2010. Section 6 concludes.

2. DEFINITION AND NOTATION

Consider a population of size n, and assume with no loss of generality that individuals receive n distinct incomes. Let $\left(y_{(1)},\ldots,y_{(i)},\ldots,y_{(n)}\right)$ be the incomes sorted in increasing order. For the aforementioned discrete distribution, the cumulative population share is (i/n), and it denotes the population share of individuals whose income is less than or equal to $y_{(i)}$. To introduce the Zenga inequality index (Zenga, 2007a; Zenga *et al.*, 2011), we define the lower mean at i as the arithmetic mean of incomes less than or equal to $y_{(i)}$

$$\bar{M}_i(Y)=\frac{1}{i}\sum_{j=1}^{i}y_{(j)}\qquad i=1,\ldots,n. \tag{1}$$

[1] The statistical literature offers works on point estimation (Langel and Tillé, 2011) and interval estimation (Greselin and Pasquazzi, 2009; Greselin, Pasquazzi and Zitikis, 2010). Polisicchio discussed the existence of continuous random variables with uniform Zenga inequality curve (2008a), and that of discrete variables with a priori fixed shape of the inequality curve (2008b). Polisicchio and Porro (2009, 2010), and Porro (2008) provided comparative studies between the Lorenz curve and the Zenga inequality curve. The behaviour of the index in empirical applications was investigated by Maffenini and Polisicchio (2010), Langel and Tillé (2009), and Zenga himself (2007b).

and the upper mean as the arithmetic mean of the remaining part of the income distribution

$$\overset{+}{M}_i(Y) = \begin{cases} \dfrac{1}{n-i}\sum_{j=i+1}^{n} y_{(j)} & i = 1, \ldots, n-1 \\ y_{(n)} & i = n \end{cases}. \tag{2}$$

Then, the Zenga point inequality measure is given by

$$I_i(Y) = \frac{\overset{+}{M}_i(Y) - \overset{-}{M}_i(Y)}{\overset{+}{M}_i(Y)}. \tag{3}$$

The inequality measure in equation (3) ranges from 0 (when the lower mean equals the upper mean) to 1 (when the lower mean equals 0). The interpretation of $I_i(Y)$ has intuitive appeal since it measures the relative income-gap between the mean income of the poorest $100(i/n)$ per cent of population and the mean income of the richest $100(1 - i/n)$ per cent of population; the smaller the disparity of the two mean incomes, the lower the value of $I_i(Y)$, thereby, implying lower inequality between the two parts of population.

Hence, the global inequality index summarizing the inequality is the average of the n inequality measures,

$$I(Y) = \frac{1}{n}\sum_{i=1}^{n} I_i(Y). \tag{4}$$

$I(Y)$ is bounded below by 0 (i.e., all individuals receive the same income) and above by $1 - 1/n^2$ (i.e., one individual receives total income).

3. DECOMPOSING INEQUALITY OVER TIME

Consider an initial time (hereafter, t_0) and a final time (henceforth, t_1) and suppose that the y individuals receive income at both times. A three-term decomposition of the inequality at t_1 is derived, showing that t_1 inequality is due to the degree of inequality at t_0, the re-ranking of individuals, and the disproportional income growth between individuals' incomes in the move from t_0 to t_1.

We start by introducing the following notation:

- Let $\left(y_{(1.0)}, \ldots, y_{(i.0)}, \ldots, y_{(n.0)}\right)$ be the t_0 incomes sorted by increasing ordering.
- Let $(y_{1.1}, \ldots, y_{i.1}, \ldots, y_{n.1})$ stand for the t_1 incomes ordered by keeping individuals lined up by the ascending ordering of their t_0 income; that is, the individual i in t_0 income parade receives the income $y_{i.1}$ at t_1.
- Let $\left(y_{(1.1)}, \ldots, y_{(i.1)}, \ldots, y_{(n.1)}\right)$ denote the t_1 incomes sorted by increasing ordering.

Re-ranking between individuals may occur in the move from the t_0 to the t_1 income distribution. Table 1 reports a simple numerical example that shows the movements of individuals within the income distribution. For instance, in Table 1 we observe that individual 1's income at t_1 (10 EUR) has rank 2 in t_1 income parade whereas his t_0 income (5.6 EUR) has rank 1 in t_0 income parade. Individual 2's income at t_0 (6.4 EUR) has rank 2 while his t_1 income (5 EUR) has rank 1 in t_1 income parade. Therefore, individual 1 and 2 exchange their positions within the income distribution in the move from t_0 to t_1.

TABLE 1. - *Re-ranking between $n = 5$ individuals*

	i	$y_{(i.0)}$	$y_{i.1}$	$y_{(i.1)}$	$I_i(Y_{(0)})$	$I_i(Y_{(1)})$
	1	5.6	10	5	0.69892	0.78947
	2	6.4	5	10	0.73529	0.73529
	3	20	25	15	0.55556	0.71429
	4	20.8	45	25	0.51471	0.69444
	5	27.2	15	45	0.41176	0.55556
Total	5	80	100	100	0.58325	0.69781

We define the income-gap $r_{i.1} = \left(y_{(i.1)} - y_{i.1}\right)$ that has to be added to individual i's income at t_1 (i.e., $y_{i.1}$) to obtain the t_1 income that has rank i in t_1 income parade (i.e., $y_{(i.1)}$),

$$y_{(i.1)} = y_{i.1} + r_{i.1}. \tag{5}$$

We then define the following quantities

$$\bar{M}_i(Y_1) = \frac{1}{i}\sum_{j=1}^{i} y_{j.1} \qquad i = 1, \dots, n \tag{6}$$

and

$$\overset{+}{M}_i(Y_1) = \begin{cases} \dfrac{1}{n-i}\displaystyle\sum_{j=i+1}^{n} y_{j.1} & i = 1, \dots, n-1 \\ y_{n.1} & i = n \end{cases}, \tag{7}$$

where $\bar{M}_i(Y_1)$ is the t_1 mean income of the poorest $100(i/n)$ per cent of population at t_0, and $\overset{+}{M}_i(Y_1)$ denotes the t_1 mean income of the richest $100(1 - i/n)$ per cent of population at t_0. Now, the mean income-gap of the poorest $100(i/n)$ per cent of population at t_0 is

$$\bar{M}_i(R_1) = \frac{1}{i}\sum_{j=1}^{i} r_{j.1} \qquad i = 1, \dots, n \tag{8}$$

and the mean income-gap of the richest $100(1 - i/n)$ per cent of the population at t_0 is

$$\overset{+}{M}_i(R_1) = \begin{cases} \dfrac{1}{n-i}\displaystyle\sum_{j=i+1}^{n} r_{j.1} & i = 1, \dots, n-1 \\ y_{(n.1)} - y_{n.1} & i = n \end{cases}. \tag{9}$$

Given equation (5) and using the properties of the arithmetic mean, one obtains

$$\overset{-}{M}_i\left(Y_{(1)}\right) = \overset{-}{M}_i(Y_1) + \overset{-}{M}_i(R_1) \tag{10}$$

and

$$\overset{+}{M}_i\left(Y_{(1)}\right) = \overset{+}{M}_i(Y_1) + \overset{+}{M}_i(R_1). \tag{11}$$

Then, the Zenga point inequality measure for $Y_{(1)}$ can be decomposed as follows,

$$\begin{aligned} I_i\left(Y_{(1)}\right) &= \frac{\overset{+}{M}_i\left(Y_{(1)}\right) - \overset{-}{M}_i\left(Y_{(1)}\right)}{\overset{+}{M}_i\left(Y_{(1)}\right)} \\ &= \frac{\left[\overset{+}{M}_i(Y_1) - \overset{-}{M}_i(Y_1)\right] + \left[\overset{+}{M}_i(R_1) - \overset{-}{M}_i(R_1)\right]}{\overset{+}{M}_i\left(Y_{(1)}\right)} \\ &= \frac{\overset{+}{M}_i(Y_1) - \overset{-}{M}_i(Y_1)}{\overset{+}{M}_i\left(Y_{(1)}\right)} + \frac{\overset{+}{M}_i(R_1) - \overset{-}{M}_i(R_1)}{\overset{+}{M}_i\left(Y_{(1)}\right)} \\ &= C_i(Y_1) + C_i(R_1). \end{aligned} \tag{12}$$

In equation (12), $C_i(R_1)$ measures the contribution to inequality at t_1 ascribable to re-ranking occurring between individuals, since in absence of re-ranking one has that $r_{i.1} = 0$ (i.e., $y_{i.1} = y_{(i.1)}$) for every $i = 1, \ldots, n$. When some individuals shift from the poorest $100(i/n)$ per cent of population at t_0 to the richest $100(1 - i/n)$ per cent of population at t_1, it follows immediately that $\overset{+}{M}_i\left(Y_{(1)}\right) > \overset{+}{M}_i(Y_1)$ and $\overset{-}{M}_i(Y_1) > \overset{-}{M}_i\left(Y_{(1)}\right)$ since $\overset{+}{M}_i(Y_1)$ is not calculated using the highest $n - i$ incomes at t_1 whereas $\overset{-}{M}_i(Y_1)$ is not the mean of the lowest i incomes at t_1. The above shown inequalities imply $C_i(R_1) > 0$, therefore, re-ranking always provides a contribution increasing inequality at t_1. From Table 2 reporting the decomposition for the example in Table 1, we can see that $C_2(R_1)$ is zero since no re-ranking occurred between any individual who was within the poorest 40 per cent of population at t_0 and any individual who was within the richest 60 per cent of population at t_0; in other words, indi-

TABLE 2. - *Calculation of the re-ranking component for $n = 5$ individuals*

	i	$y_{(i.0)}$	$y_{i.1}$	$r_{i.1}$	$y_{(i.1)}$	$C_i(Y_1)$	$C_i(R_1)$	$I_i(Y_{(1)})$
	1	5.6	10	-5	5	0.52632	0.26316	0.78947
	2	6.4	5	5	10	0.73529	0	0.73529
	3	20	25	-10	15	0.47619	0.23810	0.71429
	4	20.8	45	-20	25	-0.13889	0.83333	0.69444
	5	27.2	15	-30	45	-0.11111	0.66667	0.55556
Total	5	80	100	0	100	0.29756	0.40025	0.69781

viduals within the poorest 40 per cent of population at t_1 are the same who were within the poorest 40 per cent of population at t_0.

The interpretation of $C_i(Y_1)$ in (12) is not as straightforward as that of $C_i(R_1)$, therefore, we further decompose the former into two components. To this purpose, we define $\left(\hat{y}_{(1.1)}, \ldots, \hat{y}_{(i.1)}, \ldots, \hat{y}_{(n.1)}\right)$ that are the t_0 incomes multiplied by the scale coefficient $\lambda = \sum_{i=1}^{n} y_{(i.1)} / \sum_{i=1}^{n} y_{(i.0)}$, that is $\hat{y}_{(i.1)} = \lambda y_{(i.0)}$ for every $i = 1, \ldots, n$. This implies that $I_i\left(\hat{Y}_{(1)}\right) = I_i\left(Y_{(0)}\right)$ and $\sum_{i=1}^{n} \hat{y}_{(i.1)} = \sum_{i=1}^{n} y_{(i.1)}$.

By setting

$$d_{i.1} = y_{i.1} - \hat{y}_{(i.1)}, \tag{13}$$

equation (5) can be re-arranged as follows

$$y_{(i.1)} = \hat{y}_{(i.1)} + r_{i.1} + d_{i.1}. \tag{14}$$

In equation (14), $d_{i.1}$ is the income-gap between individual i's income at t_1 (i.e., $y_{i.1}$) and his t_0 income scaled to obtain a proportional change equal to that of the aggregate income between t_0 and t_1 (i.e., $\hat{y}_{(i.1)}$).

The mean of $d_{i.1}$ within the poorest $100(i/n)$ per cent of the population at t_0 is

$$\overset{-}{M}_i(D_1) = \frac{1}{i}\sum_{j=1}^{i} d_{j.1} \qquad i = 1, \ldots, n. \tag{15}$$

and the mean of $d_{i.1}$ within the richest $100(1 - i/n)$ per cent of population at t_0 is

$$\overset{+}{M}_i(D_1) = \begin{cases} \dfrac{1}{n-i}\displaystyle\sum_{j=i+1}^{n} d_{j.1} & i = 1, \ldots, n-1 \\ y_{n.1} - \hat{y}_{(n.1)} & i = n. \end{cases} \tag{16}$$

From equation (14) along with equations (15) and (16), $I_i\left(Y_{(1)}\right)$ can be decomposed into three terms,

$$\begin{aligned} I_i\left(Y_{(1)}\right) &= \frac{\left[\overset{+}{M}_i\left(\hat{Y}_{(1)}\right) - \overset{-}{M}_i\left(\hat{Y}_{(1)}\right)\right] + \left[\overset{+}{M}_i(R_1) - \overset{-}{M}_i(R_1)\right] + \left[\overset{+}{M}_i(D_1) - \overset{-}{M}_i(D_1)\right]}{\overset{+}{M}_i\left(Y_{(1)}\right)} \\ &= \frac{\overset{+}{M}_i\left(\hat{Y}_{(1)}\right) - \overset{-}{M}_i\left(\hat{Y}_{(1)}\right)}{\overset{+}{M}_i\left(Y_{(1)}\right)} + \frac{\overset{+}{M}_i(R_1) - \overset{-}{M}_i(R_1)}{\overset{+}{M}_i\left(Y_{(1)}\right)} + \frac{\overset{+}{M}_i(D_1) - \overset{-}{M}_i(D_1)}{\overset{+}{M}_i\left(Y_{(1)}\right)} \\ &= C_i\left(\hat{Y}_{(1)}\right) + C_i(R_1) + C_i(D_1). \end{aligned} \tag{17}$$

In equation (17), $C_i\left(\hat{Y}_{(1)}\right)$ accounts for the contribution to t_1 inequality provided by the inequality at t_0, in the sense that the distribution of $\hat{Y}_{(1)}$ preserves the ranking of individuals and the relative disparities between individuals' incomes existing at t_0; therefore, the structure of t_0 inequality is unaltered by the scale transformation applied to $Y_{(0)}$. Since $\overset{+}{M}_i(\hat{Y}_{(1)}) = \lambda \overset{+}{M}_i(Y_{(0)})$ and $\overset{-}{M}_i(\hat{Y}_{(1)}) = \lambda \overset{-}{M}_i(Y_{(0)})$, it immediately follows that $C_i(\hat{Y}_{(1)}) \geq 0$ for every $i = 1, \ldots, n$; thus, $C_i(\hat{Y}_{(1)})$ always provides a non-negative contribution to the inequality at t_1.

The component $C_i(D_1)$ in (17) measures the shift from proportional growth of aggregate income (with proportionality constant λ) shown by the poorest $100(i/n)$ per cent of population at t_0 and the richest $100(1 - i/n)$ per cent of population at t_0 when moving to the t_1 income distribution. If the richest $100(1 - i/n)$ per cent of population at t_0 enjoys proportionately larger income gain than the poorest $100(i/n)$ per cent of population at t_0, then income growth has an effect increasing inequality between the two parts of population in the move from t_0 to t_1, that is $C_i(D_1) > 0$. On the contrary, when the poorest $100(i/n)$ per cent of population at t_0 experiences proportionately larger income gain than the richest $100(1 - i/n)$ per cent of population at t_0, then income growth has an equalising effect between the two parts of population, $C_i(D_1) < 0$.

It is worth mentioning that if every individual experiences an equi-proportionate income growth equal to λ, then both $C_i(R_1)$ and $C_i(D_1)$ are equal to zero, implying $C_i(\hat{Y}_{(1)}) = I_i(Y_{(0)})$ for every $i = 1, \ldots, n$.

Table 3 reports the three-term decomposition for the numerical example shown in Table 1. We notice that $C_i(D_1)$ is zero when the poorest $100(i/n)$ per cent of the population and the richest $100(1 - i/n)$ per cent of the population at t_0 experience equi-proportionate income growth (equal to λ) in the move from t_0 to t_1 income distribution. This occurs for $i = 2$ and $i = 3$. In the remaining cases, there is disproportional income growth between the two parts of population.

For every i, $C_i(\hat{Y}_{(1)})$ is positive since it takes into account the inequality existing at t_0 (each $I_i(Y_{(0)})$ is positive as shown in Table 1).

Each point inequality measure at t_1 results from the combination of the inequality at t_0 and the concomitant effects of disproportional income growth and re-ranking. For instance, at $i = 1$ the equalizing effect of income growth (-0.15789) is more than offset by the inequality-increasing effect of re-ranking (0.26316), therefore, $I_1(Y_{(1)}) > I_1(Y_{(0)})$ (see Table 1). For $i = 2$, both re-ranking and income growth components are zero, implying that $I_2(Y_{(1)}) = C_2(\hat{Y}_{(1)}) = I_2(Y_{(0)})$.

TABLE 3. - *Decomposition of the Zenga index at t_1 for $n = 5$ individuals*

i		$\hat{y}_{i.1}$	$\hat{y}_{(i.1)}$	$d_{i.1}$	$r_{i.1}$	$y_{(i.1)}$	$C_i(\hat{Y}_{(1)})$	$C_i(D_1)$	$C_i(R_1)$	$I_i(Y_{(1)})$
	1	10	7	3	-5	5	0.68421	-0.15789	0.26316	0.78947
	2	5	8	-3	5	10	0.73529	0	0	0.73529
	3	25	25	0	-10	15	0.47619	0	0.23809	0.71427
	4	45	26	19	-20	25	0.38889	-0.52778	0.83333	0.69444
	5	15	34	-19	30	45	0.31111	-0.42222	0.66667	0.55556
Total	5	100	100	0	0	100	0.51914	-0.22158	0.40025	0.69781

The decomposition of the Zenga index at t_1 is

$$
\begin{aligned}
I\left(Y_{(1)}\right) &= \sum\nolimits_{i=1}^{n} I_i\left(Y_{(1)}\right) \\
&= \sum\nolimits_{i=1}^{n} \left[C_i\left(\hat{Y}_{(1)}\right) + C_i(R_1) + C_i(D_1)\right] \\
&= C\left(\hat{Y}_{(1)}\right) + C(R_1) + C(D_1).
\end{aligned} \tag{18}
$$

To recap, we have decomposed the t_1 inequality index into three components:

- A non-negative component accounting for the contribution provided by the t_0 inequality to the t_1 inequality, that is $C\left(\hat{Y}_{(1)}\right)$. This component is obtained by assuming that the structure of the inequality existing at t_0 is maintained unaltered at t_1.
- A component measuring the shift from proportionality shown by income growth across the t_0 income distribution, that is $C(D_1)$. When this component is negative, it means that on average income growth has an equalising effect between individuals toward the bottom of the t_0 distribution and those toward the top of the t_0 distribution. The opposite occurs when this component is positive.
- A non-negative component capturing the effect of re-ranking between individuals in the move from t_0 to t_1, that is $C(R_1)$.

4. A COMPARISON WITH THE GINI-BASED DECOMPOSITION

This section compares the proposed decomposition with the decomposition of the inequality change obtained using the generalised Gini (hereafter, S-Gini) index. Jenkins and Van Kerm (2006) showed that the link among inequality trend, re-ranking and income growth can be explained by decomposing the difference between the S-Gini index at t_1 and the S-Gini index at t_0,

$$
\begin{aligned}
\Delta G_v &= G_v\left(Y_{(1)}\right) - G_v\left(Y_{(0)}\right) \\
&= \left[1 - \sum\nolimits_{i=1}^{n} c_{i,v} L_i\left(Y_{(1)}\right)\right] - \left[1 - \sum\nolimits_{i=1}^{n} c_{i,v} L_i\left(Y_{(0)}\right)\right] \\
&= \sum\nolimits_{i=1}^{n} c_{i,v}\left[L_i\left(Y_{(0)}\right) - L_i\left(Y_{(1)}\right)\right]
\end{aligned} \tag{19}
$$

where

$$
L_i\left(Y_{(k)}\right) = \frac{1}{nM_k} \sum_{j=1}^{i} y_{(j.k)} \tag{20}
$$

is the Lorenz curve for cumulative population share (i/n) at time t_k (with $k = 0, 1$), M_k is the mean income at time t_k, and $c_{i,v}$ is

$$c_{i,v} = \begin{cases} \dfrac{(n-i+1)^v - 2(n-i) + (n-i-1)^v}{n^{v-1}} & i = 1, \ldots, n-1 \\ \dfrac{1}{n^{v-1}} & i = n \end{cases}$$

with $v \geq 1$ (Giorgi, Palmitesta and Provasi, 2006). In equation (19), v is an inequality aversion parameter which allows analysts to vary the degree of aversion to inequality. Setting $v = 2$, one has the conventional Gini index. If $v > 2$, the S-Gini index assigns poorer individuals greater social weights than the conventional Gini index does. If $1 < v < 2$, poorer individuals are assigned relatively smaller social weights compared to richer ones.

Following Jenkins and Van Kerm (2006), we introduce the concentration curve of t_1 incomes sorted by t_0 incomes, that is

$$L_i(Y_1) = \frac{1}{nM_1} \sum\nolimits_{j=1}^{i} y_{j.1}. \tag{21}$$

$L_i(Y_1)$ is calculated by keeping individuals lined up by the ascending order of their t_0 incomes. Therefore, $L_i(Y_1)$ can be interpreted as the proportion of the total income at t_1 that is received by the poorest 100(*i*/*n*) per cent of the population at t_0. By adding and subtracting $L_i(Y_1)$ in the brackets on the right hand side of equation (19), one obtains

$$\Delta G_v = R_v - P_v \tag{22}$$

where

$$R_v = \sum\nolimits_{i=1}^{n} c_{i,v} \left[L_i(Y_1) - L_i\left(Y_{(1)}\right) \right] \tag{23}$$

and

$$P_v = \sum\nolimits_{i=1}^{n} c_{i,v} \left[L_i\left(Y_1\right) - L_i\left(Y_{(0)}\right) \right]. \tag{24}$$

In equation (22) the inequality change is split into two terms: R_v accounting for re-ranking in the move from the t_0 to the t_1 distribution; P_v measuring the disproportional income growth between different parts of the income distribution (Wagstaff, 2009). R_v is an income mobility measure and it is always non-negative since $L_i(Y_1)$ cannot lay below $L_i\left(Y_{(1)}\right)$ by construction. If $P_v > 0$, individuals toward the bottom of the t_0 distribution receive proportionately larger income gains than those toward the top of the same distribution. If $P_v < 0$, income gains are proportionately larger among individuals toward the top of the t_0 distribution, implying an increase in inequality in the move from t_0 to t_1.

Comparing the decomposition shown in equation (22) with the decomposition based on the Zenga index in equation (18), we immediately notice that the latter decomposes the index at the final time instead of the difference between the indices calculated at the initial and the final time. The longitudinal decomposition of the

Zenga index directly explains the link between the initial inequality and the re-ranking and income growth effects, since it clearly detects the roles played by the initial inequality, re-ranking and income growth in determining the inequality at the final time.

A further difference between the Gini-based approach and the Zenga-based approach is that the former incorporates an inequality aversion parameter in the decomposition (i.e., v), allowing analysts to vary the degree of aversion to inequality. This is an appreciable feature of the Gini-based approach. Future research may address the generalisation of the Zenga index formulation in order to account for the degree of inequality aversion.

Our last remark concerns the problem of right censoring when observing trends in inequality over time. When releasing public use income data, risks of disclosure are limited by replacing income values which are above an income threshold by the threshold itself (“top code”). As pointed out in Jenkins, Burkhause, Feng and Larrimore (2011), if top code values are not adjusted consistently over time, top coding can affect the estimates of inequality measures and their standard errors. To account for the uncertainty added to estimates by the use of top-coded observations, Jenkins *et al*. (2011) introduced statistics to test the significance of the difference between inequality measures calculated at different times. Even though our decomposition approach does not rely on the difference between inequality indices to explain the inequality trend, we argue that top-coded observations inevitably affect the estimate of the Zenga index at t_1.

5. EMPIRICAL APPLICATION

We applied the decomposition to household incomes collected by the Survey on Household Income and Wealth (henceforth, SHIW) of the Bank of Italy in 2008 and 2010 (Banca d’Italia, 2012). The SHIW is carried out every two years, and each survey sample comprises households interviewed for the first time and households interviewed in previous surveys (panel households). The sample size of the 2010 survey is 7,951, including 4,621 panel households interviewed in the 2008 survey. We performed the analysis using 2008 and 2010 incomes for these panel households. We excluded panel households with income equal to or less than zero in either 2008 or 2010. The retained households were 4,611. The mean income is 33,160.8 EUR in 2008 and 33,245.8 in 2010. Table 4 shows the quintiles of the 2008 and the 2010 income distributions. From Table 4, we note that the distribution of total income among the quintiles is almost unaltered between the two years. However, households within the various quintiles are not necessarily the same in both years, since some households may have shifted from a quintile to another in the move from 2008 to 2010.

Figure 1 plots the Zenga inequality curve for the 2008 and the 2010 income distributions. The 2010 inequality curve lies above the 2008 one for lower values of p whereas it is below the 2008 inequality curve when p increases, however, since the

TABLE 4. - *Quintiles of income distribution, Italy, 2008-2010*

2008		2010	
Income class	% of total income	Income class	% of total income
0 ┤ 16,558.6	3.25	0 ┤ 16,622.8	3.22
16,558.6 ┤ 23,647.5	8.71	16,622.8 ┤ 23,973.1	8.72
23,647.5 ┤ 32,285.4	16.28	23,973.1 ┤ 32,630	16.32
32,285.4 ┤ 45,087	26.69	32,630 ┤ 45,908.9	26.81
45,087 ┤ 395,120.2	45.07	45,908.9 ┤ 292,258.7	44.93
Total	100		100

Source: Bank of Italy, *calculations on SHIW 2008 and 2010 data.*

two curves intersects we cannot unambiguously say whether overall inequality decreases or increases over the 2008-2010 period.

The global inequality indices as well as the decomposition of 2010 index are reported in Table 5. From Table 5, we observe that inequality slightly decreases between 2008 and 2010. This finding is in line with the very small reduction in inequality observed between the Gini indices calculated in 2008 and 2010 using the whole sample data of the two surveys (Banca d'Italia, 2012, p. 15); therefore, the almost stable inequality over the 2008-2010 period is confirmed even though we restricted our attention to panel households.

The decomposition results indicate that the inequality existing in 2008 plays the most important role in determining the inequality in 2010, since the contribution of 2008 inequality (0.681267) is greater than the inequality index in 2010 (0.672693). This outcome is not surprising once it has been observed that the inequality in 2008 (0.675291) is higher than the inequality in 2010 (0.672693).

TABLE 5. - *Decomposition of the Zenga index over time, Italy, 2008-2010*

$I(Y_{(08)})$	$I(Y_{(10)})$	$C(\hat{Y}_{(10)})$	$C(D_{10})$	$C(R_{10})$
0.675291	0.672693	0.681267	-0.105356	0.096783

Source: Bank of Italy, *calculations on SHIW 2008 and 2010 data.*

Even though the values of $I(Y_{(08)})$ and $I(Y_{(10)})$ are close, the decomposition shows that the inequality trend is affected by the households' movements and the disproportional income growth between households' incomes. The income growth produces an equalising effect (-0.105356) on inequality in the move from 2008 to 2010. This effect reducing inequality is partially offset by the re-ranking contribution (0.096783) which increases inequality.

We now turn to the analysis of the decomposition of the Zenga inequality curve in Figure 2, where every component is expressed as a percentage of the Zenga inequality curve for the 2010 income distribution.

From Figure 2, we observe that the income growth component always lies below zero, therefore, income growth benefits the poorest 100p per cent of the population

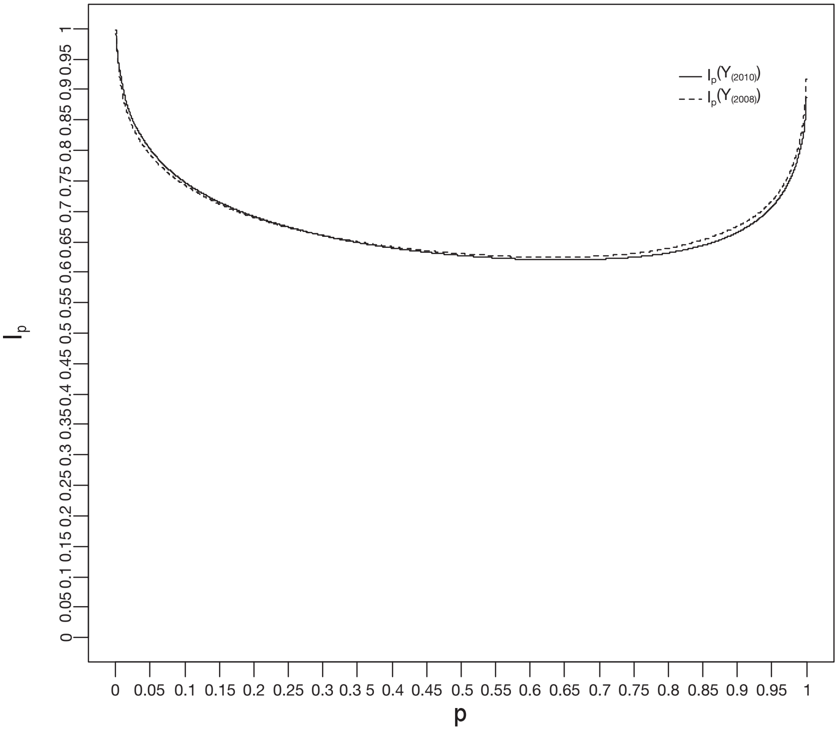

FIGURE 1. - *Zenga's inequality curves, Italy, 2008-2010*

in 2008 proportionally more than the richest 100(1-*p*) per cent for every *p*. This unambiguously indicates that income growth has an equalising effect on incomes in the move from 2008 to 2010, as summarised by $C(D_{10})$.

The re-ranking component provides a positive contribution to inequality that is greater than 11.8 per cent of the point inequality measure at every *p*. Thus, given any *p*, there is at least a pair of households (one belonging to the poorest 100*p* per cent of the population in 2008, the other belonging to the richest 100(1-*p*) per cent) whose members re-ranked in the move from 2008 to 2010.

Figure 2 shows that re-ranking and income growth effects on 2010 inequality are larger at the bottom and the top of the income distribution. The income growth component is more than offset by the re-ranking one at the bottom of the distribution, whereas the opposite occurs at the top. More specifically, the increase in inequality at the bottom of the income distribution between 2008 and 2010 (see Figure 1) is explained by observing that $C_p(R_{10})$ slightly exceeds $C_p(D_{10})$ in absolute value for *p* ranging from 0.0026 to 0.2889. On the contrary, the impact of disproportional income growth at the top of the distribution is large enough to counterbalance the re-ranking effect and to reduce the inequality existing in 2008.

Overall, our findings show that both re-ranking and income growth affect the inequality change between the two years, even though the inequality change is very

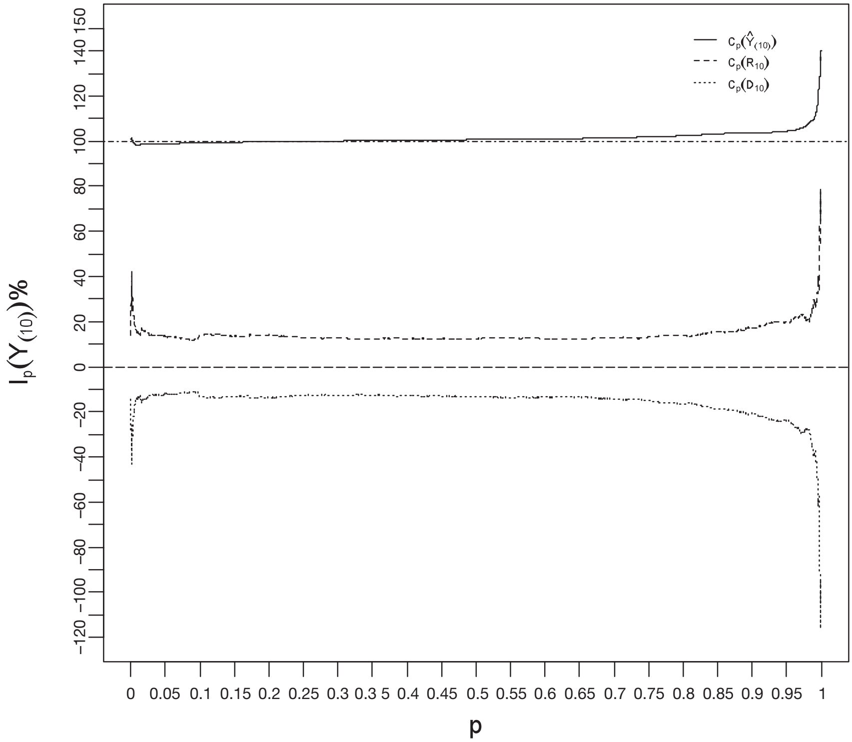

FIGURE 2. - *Decomposition of Zenga's inequality curve for the 2010 income distribution, Italy*

small; that is, the nearly stable inequality is not simply due to few changes in the income distribution. This emphasises that observing only the inequality trend can lead to misleading conclusions, since an almost unaltered inequality may conceal that income receivers change their fortunes and exchange their respective positions over time.

6. CONCLUSIONS

In this study, we developed a decomposition of the Zenga index that explains the link among inequality trend, re-ranking, and income growth. We showed that inequality at a certain time is given by the interaction between re-ranking and income growth effects which modify the inequality existing at an earlier time. This decomposition approach provides new insights for the analysis of inequality trend, since the individuals' movements and the disproportional growth of their incomes cannot be taken into account by referring solely to the change in inequality occurring between two points in time.

The decomposition was applied to Italian household income over the 2008-2010 period. Quite unsurprisingly, large part of the 2010 inequality is due to the 2008 inequality. The very small reduction in income inequality over the 2008-2010 period is explained by the income growth effect which offsets the disequalising effect of re-ranking. This suggests that the nearly stable inequality is the outcome of the interaction between re-ranking and disproportional income growth rather than an almost equi-proportionate income growth.

Future research will address the decomposition when the time elapsed between the initial and the final year is larger than two years. Moreover, we intend to investigate the possibility of intersecting the longitudinal dimension of decomposition with the source and subgroup decomposition dimensions.

REFERENCES

Banca d'Italia (2012). Survey on Household Income and Wealth 2010. Available at http://www.bancaditalia.it/statistiche/indcamp/bilfait/boll_stat.

Giorgi G.M., Palmitesta P., Provasi C. (2006). Asymptotic and bootstrap inference for the generalized Gini indices. *Metron*, **64**, 107-124.

Greselin F., Pasquazzi L. (2009). Asymptotic Confidence Intervals for a New Inequality Measure. *Communications in Statistics - Simulation and Computation*, **38**, 1742-1756.

Greselin F., Pasquazzi L., Zitikis R. (2010). Zenga's new index of economic inequality, its estimation and an analysis of incomes in Italy. *Journal of Probabability and Statistics*, **2010** (2010): Article ID 718905, 26 p, doi:10.1155/2010/718905.

Jenkins S.P., Burkhauser R.V., Feng S., Larrimore J. (2011). Measuring inequality using censored data: a multiple-imputation approach to estimation and inference. *Journal of the Royal Statistical Society A*, **174**, 63-81.

Jenkins S.P., Van Kerm P. (2006). Trend in income inequality, pro-poor income growth, and income mobility. *Oxford Economic Papers*, **58**, 531-548.

Langel M., Tillé Y. (2009). An evaluation of the performance of inequality measures for the detection of changes in an income distribution. Technical report, University of Neuchatel.

Langel M., Tillé Y. (2012). Inference by linearization for Zenga's new inequality index: a comparison with the Gini index. *Metrika*, **75**, 1093-1110.

Maffenini W., Polisicchio M. (2010). How potential is the $I_{(p)}$ inequality curve in the analysis of empirical distributions. DiMeQUANT Working Paper n. 186, University of Milan Bicocca.

O'Neill D., Van Kerm P. (2008). An integrated framework for analysing income convergence. *The Manchester School*, **76**, 1-20.

Polisicchio M. (2008a). The continuous random variable with uniform point inequality measure I(p). *Statistica & Applicazioni*, **VI**, 137-151.

Polisicchio M., Porro F. (2010). How the parameters of the new Zenga distribution model affect the inequality curves $L_{(p)}$ and $I_{(p)}$. Annual meeting of the German Statistical Society, Nuremberg, September 14-17, 2010.

Polisicchio M., Porro F. (2009) A comparison between Lorenz $L_{(p)}$ curve and Zenga $I_{(p)}$ curve. *Statistica Applicata*, **21**, 289-301.

Porro F. (2008) Equivalence between partial order based on $L_{(p)}$curve and partial order based on $I_{(p)}$curve. Proceeding of the XLIV Scientific Meeting of the Italian Statistical Society, 25-27 June 2008, Arcavacata di Rende.

Radaelli P. (2008). A subgroups decomposition of Zenga's uniformity and inequality indexes. *Statistica & Applicazioni*, **VI**, 117-136.

Radaelli P. (2010). On the decomposition by subgroups of the Gini index and Zenga's uniformity and inequality indexes. *International Statistical Review*, **78**, 81-101.

Wagstaff A. (2009). Reranking and Pro-Poor Growth: Decompositions for China and Vietnam. *Journal of Development Studies*, **45**(9), 1403-1425.

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

Zenga M.M. (2007b). Applications of a new inequality curve and inequality index based on ratios between lower and upper arithmetic means. In M.I. Gomes, D. Pestana, P. Silva (eds.), *Bulletin of the International Statistical Institute*, volume Proceedings on CD, 56th World Statistics Congress of the International Statistical Institute. Lisbon, Portugal.

Zenga M.M., Radaelli P., Zenga Ma. (2011). Decomposition of Zenga's inequality index by sources. DiMeQUANT Working Paper n. 216, University of Milan Bicocca.