Gini versus Zenga Inequality Index: a New Approach to Measuring
Personal Income Tax Redistribution and Progressivity

March 7th, 2017 --- 12.31

Francesca Greselin
University of Milan Bicocca (IT)
Phone: +39 02 64 483 118
E-mail: francesca.greselin@unimib.it

Simone Pellegrino (Corresponding Author)
University of Turin (IT)
Phone: +39 011 670 6060
E-mail: simone.pellegrino@unito.it

Achille Vernizzi
University of Milan (IT)
Phone: +39 02 50 321 460
E-mail: achille.vernizzi@unimi.it

Abstract
In this paper we introduce a new methodology to study the degree of progression, the redistributive and re-ranking effects of a personal income tax system, by employing and extending the new inequality curve and index proposed by Zenga (2007). Given an income distribution, the Zenga curve compares the economic conditions of two exhaustive groups of population obtained by dividing the overall population at all possible percentiles, from the bottom to the top observed income. Since the recent literature underlines that the Zenga curve shows features that are different from the standard approach based on the Lorenz curves, we show the potentialities of the new curve (and index) when studying the effects exerted by a personal income tax. This new methodology is compared to the classical one by a stylized example and by developing an application to Italian personal income tax data.

JEL Codes: H23, H24

Keywords: Personal Income Tax, Gini Index, Micro-simulation Models, Reynolds-Smolensky Index, Kakwani index, Zenga Index
1. Introduction

The economic literature, in a long journey of more than a hundred years, has proposed several approaches to studying the inequality of quantitative variables, primarily income distributions. These approaches can be categorized in two basic groups: the first one is aimed at providing graphical representations of inequality (i.e., by plotting the frequency and cumulative density functions, the Pen’s parade, and via the most popular Lorenz and concentration curves); the second group comprises positive and normative measures of inequality (i.e., variance, generalized entropy index, Theil index, Atkinson index, Gini coefficient, Lorenz curve).

Synthetic indexes allow the inequality of income distributions to be summarized and compared by means of a single scalar. Among these, the most famous inequality index is undoubtedly the Gini coefficient, which has also a graphical explanation through the Lorenz curve. Keeping in mind that relative income differentials get compressed in the transition from the pre- to the post-tax distribution, the tax literature has proposed a Gini-based methodology, to measure the degree of progression of taxation, using the Kakwani index, as well as the redistributive effect exerted by the tax via the Reynolds-Smolensky index.

Whilst the Lorenz curve is a fundamental tool for welfare comparisons (Atkinson, 1970; Shorrocks, 1983; Dardanoni and Lambert, 2001), its application in the tax literature is overall poor, since generally it is not easy to draw specific conclusions by examining and comparing Lorenz and concentration curves of different distributions (i.e., pre-tax distribution, post-tax distribution and tax distribution). As a consequence, in most of the existing literature the overall effect of taxation is primarily derived from the Gini and concentration coefficients.

Recently, Zenga (2007) proposed a new methodology to both plot and measure inequality. In this paper we will show that the graphical representation provided by the Zenga inequality curve is an accurate instrument for understanding which part of a pre-tax distribution is mostly affected by the tax. Basically, the curve and the index introduced by Zenga are based on a comparison of the mean income of the poorer income earners with the mean income of the remaining richest part of the population. This methodology gives more complete information than the usual Lorenz-based
approach, so that in this paper we extend the Zenga methodology to the study of the redistributive effect as well as the degree of progression exerted by the personal income tax, showing its strengths and weaknesses.

The paper is organized as follows. Section 2 describes the standard Lorenz and Gini approaches, while Section 3 turns to the new Zenga curve and index. Section 4 presents formulas for the computation of the Lorenz and Zenga curves using real data. To analyze the effects of taxation on income distributions, Section 5 extends the Zenga curve and index, introducing new tools for measuring the degree of progression of a personal income as well as its re-ranking and redistributive effects. In Section 6 an explanatory and stylized example is discussed, to show how the Zenga curve and index can be employed to draw conclusions when considering tax reform. Section 7 presents the microsimulation model employed for the empirical estimations, particularly useful for developing our analysis of the Italian personal income tax (Section 8). Section 9 makes some concluding remarks.

2. The Lorenz curve and the Gini index

Given a random variable \( Z \geq 0 \) with non negatively supported \( cdf F(Z) \) for \( Z \geq 0 \) representing gross or net incomes as well as taxes, we denote by \( F^{-1}(p) = \inf \{ z : F(z) \geq p \} \) the corresponding population quantile function for \( 0 < p < 1 \). The notion of the Lorenz curve was introduced by M.O. Lorenz in 1905 to plot the cumulative share of \( Z \), denoted by \( L_F(p) \), versus the cumulative share of the population \( p \). In the ideal case of perfect equality (that is, a society in which all people have the same income) the share of incomes equals the share of the population, so that \( L_F(p) = p \), for all \( 0 < p < 1 \). In this case the Lorenz curve is the diagonal line from \((0,0)\) to \((1,1)\). On the other hand, the lower the share of income \( L_F(p) \) held by the share of income earners \( p \), the higher the inequality. In the ideal case of perfect inequality (that is, a society in which all people but one have an income of nil) the share of
incomes equals zero for $0 \leq p < 1$, so that $L_F(p) = 0$, and only for $p=1$ we have $L_F(1) = 1$.

Hence, the standard Lorenz curve is given by $(p, L_F(p))$, where

$$L_F(p) = \frac{\int_0^p F^{-1}(s)ds}{\int_0^1 F^{-1}(s)ds} = \frac{1}{\mu_F} \int_0^p F^{-1}(s)ds,$$  \hspace{1cm} (1)

and $\mu_F = \mu(Z)$ denotes the mean value, or the expectation of the random variable $Z$.

It seems very natural to express the degree of inequality through the deviation of the actual Lorenz curve from the diagonal line.

The Gini index, introduced by C. Gini in 1914, is precisely given by twice the area between the equality line and the Lorenz curve, as follows

$$G_F = 2\int_0^1 (p - L_F(p)) dp. \hspace{1cm} (2)$$

From a historical point of view, it is interesting to recall that Gini moved from proposing the so-called Gini inequality curve

$$G_F(p) = \frac{\mu_F - \frac{1}{p_0} \int_0^p F^{-1}(s)ds}{\mu_F},$$  \hspace{1cm} (3)

that expresses the relative deviation of the mean income of the poorer $p\%$ of the population from the overall mean. Observing that

$$\frac{\mu_F - \frac{1}{p_0} \int_0^p F^{-1}(s)ds}{\mu_F} = \frac{p - L_F(p)}{p},$$  \hspace{1cm} (4)

we see that the Gini index is the weighted mean of the Gini inequality curve, with weights given by $2p$:

$$G_F = 2\int_0^1 \frac{(p - L_F(p))}{p} dp. \hspace{1cm} (5)$$
3. The new Zenga curve and index

Recently, Zenga (2007) proposed a different approach for measuring and representing inequality. Observing the noticeable increase in disparities between less fortunate and more fortunate individuals, Zenga introduced a new inequality curve \( I_p(p) \) obtained by contrasting the average income of the poorer \( p\% \) bottom earners \( \frac{1}{p} \int_0^{\frac{p}{1-p}} F^{-1}(s) \, ds \) with the amount that is held, on average, by the richest top earners, i.e. the remaining \((1-p)\%\) of the population, that is \( \frac{1}{1-p} \int_{\frac{p}{1-p}}^{1} F^{-1}(s) \, ds \).

Therefore, he defined the curve \((p, I_p(p))\) where

\[
I_p(p) = \frac{1}{1-p} \int_{\frac{p}{1-p}}^{1} F^{-1}(s) \, ds - \frac{1}{p} \int_0^{\frac{p}{1-p}} F^{-1}(s) \, ds, \quad \text{for} \quad 0 < p < 1.
\]  

We see that, when the random variable \( Z \) is equal to a constant, the corresponding quantile \( F^{-1}(p) \) is also equal to the constant and thus \( I_p(p) = 0 \) for every \( p \in (0,1) \), meaning perfect equality or the “egalitarian society.” The other extreme scenario is when, loosely speaking, there is only one member in the society who gets the entire income of the population, and in this case \( I_p(p) = 1 \) for every \( p \in (0,1) \).

As illustrated by Greselin et al. (2010), the methodology proposed by Zenga (2007) keeps “in mind that the notions of poor and rich are relative to each other” and summarizes, in a single measure, the amount of inequality in the population, by proposing the following summary index

\[
I_F = \int_0^1 I_p(p) \, dp.
\]

Although the literature on the Zenga index and curve is not (obviously) as copious as that on the Gini index, we find research on various properties of the index and curve (Polisicchio 2008, Polisicchio and Porro, 2009; Maffènini and Polisicchio, 2012; Greselin, Puri and Zitikis 2009; Arcagni and Porro, 2014), inferential theories and their applications (Greselin and Pasquazzi, 2009; Greselin, Pasquazzi and Zitikis 2010, 2013,
subgroup decompositions of the index (Radaelli 2008, 2010), a longitudinal decomposition (Zenga and Mussini 2010) and decompositions by income sources (Zenga et al 2012; Zenga, 2013), as well as applications on real data (Arcagni and Zenga, 2013). We also find detailed discussions of the advantages of the Zenga index over the Gini index within both the descriptive and inferential frameworks. Langel and Tillé (2012) have revealed certain characteristics of the sampling distribution of the empirical Zenga index, facilitating more reliable inferential results. Antal et al. (2012) have extended inferential results to complex sampling designs.

Comparing now the Lorenz and the Zenga curves, we observe that the Zenga one neither has ‘forced’ values at the end-points of its domain of definition nor is it constrained to be non-decreasing and concave on the interval $[0,1]$, as is the case with the Lorenz curve. Moreover, while $I_F(p)$ compares the mean incomes of two disjointed sub-populations, the poor and the rich, $G_F(p)$ and $L_F(p)$ compare overlapping parts of the population.

The last consideration is related to the weight function $2p$, adopted to obtain the Gini index for normalization purposes, as in (5). It gives greater emphasis to the comparisons of almost coinciding sub-populations, which are likely to be less informative, while the Zenga index takes into account, with the same weight, any relative deviation from inequality, measured by $I_F(p)$, in any part of the distribution.

### 4. Empirical estimates of the Gini and Zenga indexes

In the previous section, we have presented the Gini and the Zenga curves and indexes as functionals, defined on the space of the distribution functions. In this section we address ourselves to how they are to be estimated on real data. Let us now suppose that there are $n$ households in a real sample, and let $z_1, \ldots, z_n$ be the observed ranked values of a quantitative variable $Z$ (say gross income, tax or net income).
We have seen that the Zenga index compares the average of the attribute at stake from the first household to the \( i \)-th one, say \( M_i(Z) = \frac{1}{i} \sum_{j=1}^{i} z_j \), with the average of the remaining \( n-i \) households \( M_i'(Z) = \frac{1}{n-i} \sum_{j=i+1}^{n} z_j \). Therefore, the ratio

\[
I_i(Z) = \frac{M_i(Z) - M_i'(Z)}{M_i(Z)}
\]  

(8)
describes the inequality at each percentile of the distribution. If all income earners hold the same income, all ratios are equal; conversely their variability is a direct function of inequality.

We will consider also the complementary curve \( U_i(Z) \), called the *uniformity* curve, defined by

\[
U_i(Z) = 1 - I_i(Z) = \frac{M_i'(Z)}{M_i'(Z)}
\]  

(9)

which has the advantage of a more straightforward interpretation, since it simply measures \( M_i(Z) \) in terms of a percentage of \( M_i'(Z) \). Both \( U_i(Z) \) and \( I_i(Z) \) potentially range between 0 and 1 (Zenga, 2007), and are not constrained in their end points, for \( i=1/n \) and \( i=(n-1)/n \), differently from \( L_i(Z) = \frac{\sum_{j=1}^{i} z_j}{\sum_{j=1}^{n} z_i} \), that always begins in (0,0) and ends in (1,1).

Moreover, \( I_i(Z) \) and \( U_i(Z) \) can be expressed in terms of \( L_i(Z) \) (see Zenga, 2007):

\[
I_i(Z) = 1 - \frac{p_i}{p_i} \frac{L_i(Z)}{1-L_i(Z)}
\]  

(10)

\[
U_i(Z) = 1 - \frac{1-p_i}{p_i} \frac{L_i(Z)}{1-L_i(Z)}
\]  

(11)

Finally, we arrive at the empirical estimators \( \hat{I}(Z) \) of the Zenga index (Greselin et al., 2010):

\[
\hat{I}(Z) = \frac{1}{n-i} \sum_{j=i+1}^{n} \frac{1}{n-i} \sum_{j=1}^{i} z_j - \frac{1}{i} \sum_{j=1}^{n} z_j
\]  

(12)
and \( \hat{G}(Z) \) for the Gini index:

\[
\hat{G}(Z) = -\frac{1}{n-1} \sum_{j=1}^{n} \frac{1}{n} \sum_{j=1}^{n} z_j - \frac{1}{i} \sum_{j=1}^{i} z_j - \frac{1}{n} \sum_{j=1}^{n} z_j - \frac{1}{2i}. \tag{13}
\]

5. The analysis of the redistributive effect of taxes by the Zenga approach

In the following we jointly analyze the pre- and post-tax income distributions as well as the tax distribution. First, we present the standard approach based on the Lorenz curves and the Gini coefficient, and then we show how analogous curves and indexes can be defined on the basis of the Zenga approach.

Supposing that the pre-tax income distribution \( X \) represents incomes in non-decreasing order, non necessarily the paired samples of the after-tax incomes \( Y \) and taxes \( T \) are still similarly ranked, so that for each observed pair of values \( x_i < x_j \) it is not granted that \( y_i < y_j \) and/or \( t_i < t_j \). As a consequence, we denote by \( Y \) the post-tax income distribution when net incomes are ordered in non-decreasing order, and by \( Y_X \) the after-tax income distribution when income units are ranked according to the pre-tax order. Similarly, we call \( T \) the tax distribution when taxes are ordered in non-decreasing order, and denote by \( T_X \) the tax distribution when the income units related to \( T \) remain ranked according to the pre-tax order.

To evaluate the redistributive effect and the degree of progression, as well as the re-ranking effect of a tax system, the specialized literature has produced several indexes, which are mainly functions of the Gini coefficients, \( G(\cdot) \), of \( X \) and \( Y \) and of the concentration indexes, \( C(\cdot) \), of \( Y_X \) and \( T_X \).

A global measure of tax progressivity assesses the deviation of a given tax system from proportionality, hence it is related to the local index of liability progression, that is, the elasticity of the tax liability with respect to the pre-tax income evaluated at each pre-tax income level. The overall degree of progression is generally evaluated by the Kakwani (1976) index

\[
K = C(T_X) - G(X). \tag{14}
\]
which can be represented as twice the area between the Lorenz curve of $X$, and the concentration curve of $T_X$.

Following the same strategy, we here introduce an analogous curve and synthetic measure, in accordance with the Zenga approach

$$KI_i = I_i(T_X) - I_i(X) = U_i(X) - U_i(T_X),$$

$$KI = I(T_X) - I(X) = U(X) - U(T_X).$$

We see that $KI_i$ involves differences between the tax uniformity curve, $U_i(T_X)$, and the pre-tax uniformity Zenga curve, $U_i(X)$.

We know that if the concentration of taxes is greater than the concentration of pre-tax incomes, the post-tax income distribution is less concentrated than the pre-tax one, and the tax is progressive. If we compare the Gini coefficient of the pre-tax distribution with the post-tax concentration, maintaining the ordering of income units according to the pre-tax income rank, we yield the Reynolds-Smolensky index $RS$; if we compare the Gini coefficients of the pre-tax and the post-tax income distribution, respectively, we yield the $RE$ index. Hence, the redistributive effect $RE$ is usually quantified as twice the area between the Lorenz curves for pre- and post-tax distributions; and the Reynolds-Smolensky index $RS$, given by twice the area between the Lorenz curve for the pre-tax distribution and the concentration curve for the post-tax distribution.\footnote{Reynolds-Smolensky (1977); see e.g. Lambert (2001), page 207.}

Formally:

$$RE = G(X) - G(Y) = [G(X) - C(Y_X)] - [G(Y) - C(Y_X)],$$

$$RS = G(X) - C(Y_X).$$

The Reynolds-Smolensky $RS$ index compares the Gini coefficient for the pre-tax income distribution with the concentration index for the after-tax income distribution, that is the concentration evaluated when income units are sorted according to the pre-tax ordering; $RE$ is given by the difference between the Gini coefficients of pre- and post-tax distributions. It is immediately observed that $RS \geq RE$; the strict inequality holds where the tax determines re-ranking, so that $R(Y_X) > 0$, when passing from pre- to post-
tax income distribution; the last index $G(Y) - C(Y_X)$ is a measure of the overall re-ranking, also known as the Atkinson, Plotnick, Kakwani index.\(^2\)

Following the Zenga approach we are able to introduce the new curves and synthetic indexes, respectively:

\[
REI_i = I_i(X) - I_i(Y) = I_i(X) - I_i(Y_X) - RI_i(Y_X) = U_i(Y_X) - U_i(X) - RI(Y_X),
\]

\[
REI = I(X) - I(Y) = I(X) - I(Y_X) - RI(Y_X) = U(Y_X) - U(X) - RI(Y_X),
\]

\[
RS_i = I_i(X) - I_i(Y_X) = U_i(Y_X) - U_i(X),
\]

\[
RS = I(X) - I(Y_X) = U(Y_X) - U(X),
\]

\[
RI_i = I_i(Y) - I_i(Y_X) = U_i(Y_X) - U_i(Y).
\]

It is well known that the Kakwani progressivity index $K$ is related to the Reynolds-Smolensky $RS$ index by the relation (Lambert, 2001):\(^3\)

\[
RS = [M(T)/M(Y)]K,
\]

where $M(T)$ and $M(Y)$ are the averages of $T$ and $Y$, respectively. This means that $RS$ is a function of two variables, the Kakwani index $K$ and the overall average tax rate; therefore $RS$ can increase even if the overall average tax rate decreases, if $K$ more than compensates the effect due to the tax rate.

It is straightforward to show that $RS_i$ and $KI$ are linked by an analogous relation:

\[
RS_i = I_i(X) - I_i(Y_X) = \hat{\lambda}_i [I_i(T_X) - I_i(X)] = \hat{\lambda}_i \cdot KI_i.
\]

To verify (25), let us start from the following equality

\[
\frac{M_i^+(X) - M_i^-(X)}{M_i^-(X)} - \frac{M_i^+(Y_X) - M_i^-(Y_X)}{M_i^-(X)} = \frac{M_i^+(T_X) - M_i^-(T_X)}{M_i^+(X)}
\]

which is immediately verified. From the latter equation we get:

\[
\left[ \frac{M_i^+(X) - M_i^-(X)}{M_i^+(X)} - \frac{M_i^+(Y_X) - M_i^-(Y_X)}{M_i^+(X)} \right] \frac{M_i^+(X)}{M_i^+(T_X)} = \frac{M_i^+(T_X) - M_i^-(T_X)}{M_i^+(X)} \frac{M_i^+(T_X)}{M_i^+(Y_X)}.
\]

which, after trivial simplifications, becomes

\(^2\) See e.g. Lambert (2001), page 241.

\(^3\) See e.g. Lambert (2001), page 241.
\[
\frac{M_i^+(X) - M_i^-(X)}{M_i^+(Y_X)} - \frac{M_i^+(Y_X) - M_i^-(Y_X)}{M_i^+(Y_X)} = \frac{M_i^+(T_X) - M_i^-(T_X)}{M_i^+(T_X)} \frac{M_i^+(T_X)}{M_i^+(Y_X)}.
\]

If we add to the l.h.s. of the previous equation the quantity

\[
\frac{M_i^+(X) - M_i^-(X)}{M_i^+(X)} - \frac{M_i^+(X) - M_i^-(X)}{M_i^+(Y_X)}
\]

and to the r.h.s. the equivalent quantity

\[
\frac{M_i^+(X) - M_i^-(X)}{M_i^+(X)} - \frac{M_i^+(X) - M_i^-(X)}{M_i^+(Y_X)}
\]

it yields

\[
\frac{M_i^+(X) - M_i^-(X)}{M_i^+(X)} - \frac{M_i^+(X) - M_i^-(X)}{M_i^+(Y_X)}
\]

which is expression (26), with \( \lambda_i = \frac{M_i^+(T_X)\lambda_i}{M_i^+(Y_X)} \).

In equation (26) \( I_i(T_X) - I_i(X) \) measures the tax progressivity, at point \( i \), whilst the factor \( \lambda_i = \frac{M_i^+(T_X)}{M_i^+(Y_X)} \) measures the tax incidence for incomes greater than \( x_i \); if the tax is progressive, \( \lambda \) is an increasing function of \( X \), so that the sequence \( (\lambda_1, \lambda_2, \ldots, \lambda_{n-1}) \) reflects the tax system progressivity.

The synthetic version of (26) can be expressed as follows:

\[
RSI = I(X) - I(Y_X) = [I(T_X) - I(X)] \lambda^*.
\]

where

\[
\lambda^* = \frac{1}{n-1} \sum_{i=1}^{n-1} \frac{I_i(T_X) - I_i(X)}{[I(X) - I(T_X)]} \lambda_i.
\]

By using (26) and (27), the curve and synthetic REI index can be re-written, respectively, as

\[
REI_i = I_i(X) - I_i(Y) = [I_i(T_X) - I_i(X)] \lambda_i - RI(X),
\]

\[
REI = I(X) - I(Y) = [I(T_X) - I(X)] \lambda^* - RI(X).
\]
In the curve (equation 26), the factor which multiplies \( I_i(T_X) - I_i(X) \) is generally different for each point \( i \), as well as for the synthetic term \( I(T_X) - I(X) \). The same does not apply in the Gini based approach: the difference between the ordinates of the concentration curve of \( T_X \) and the corresponding ones of the Lorenz curve of \( X \), is multiplied in all cases by the factor \( M(T)/M(Y) \), and the same factor multiplies the synthetic difference \( C(T_X) - G(X) \) as well.

6. Evaluating a personal income tax reform: a theoretical example

To illustrate the curves and indices that we have discussed so far, we provide a stylized example using hypothetical data. Consider a lognormal distribution of 10 thousand pre-tax incomes \( X \), with the mean value equal to €20,138, and the lowest and the highest observed values equal to €431 and €310,350, respectively; suppose that the weight is equal to 1 for all incomes. A rate schedule is applied to the gross income: 10 percent for gross incomes below €10,000 (which corresponds to the 36th percentile); 30 percent for gross incomes between €10,000 and €40,000 (which corresponds to the 91st percentile); 50 percent for gross incomes above €40,000 (tax system 1).

No deductions and tax credits are allowed, so that no re-ranking can occur in the transition from the pre- to the post-tax incomes. Let us consider a tax reform augmenting the first marginal tax rate from 10 to 20 percent (tax system 2). This tax reform is obviously regressive. All taxpayers face a higher tax liability, and, in particular, two cases emerge: taxpayers belonging to the bottom 36 percentiles (gross income equal to or below €10,000) double their tax liability; for all the other taxpayers, the tax increase is exactly €1,000. In both tax systems the average tax rates are not decreasing and their difference decreases with income.

The Lorenz curves for tax liability before and after the reform, denoted by \( L_i(T^1) \) and \( L_i(T^2) \), tell us only a part of the story (Figure 4). They emphasize that the share of the overall tax increase increases with income: the bottom decile pays 3.8 percent of the
overall tax increase, the second one pays 6.7 percent, while the third pays 9.1 percent; each of the remaining deciles pays 11.5 percent of the overall tax increase.

**Figure 4: Lorenz curves**

![Lorenz curves](image)

Within this framework, information regarding two contrasting aspects is missed: on the one hand, the poorest 36 percent of taxpayers double their tax liability, which is a lot; on the other hand, their contribution to the overall tax increase is low, since they are characterized by very low incomes. The Lorenz curves for tax liability do not make these two facts apparent.

By contrast, the two Zenga uniformity curves for tax liability \( U(T^1) \) and \( U(T^2) \) are further away from each other and show an unexpected deviation from concavity that can help us to understand which part of the pre-tax distribution is affected the most by this tax reform (Figure 5). Let us consider the solid red curve \( U(T^1) \), that is the Zenga curve for tax liability before the tax reform (tax system 1). To see what information is conveyed by such a curve, let us say that the point (0.1, 0.064) tells us that the average tax liability of taxpayers belonging to the bottom 10 percent of the distribution is just 6.4 percent of the average tax liability of the remaining 90 percent of taxpayers. The slope of this curve increases, starting from the 36th percentile, that is, starting from incomes belonging to the second tax bracket. The reason is manifest: the marginal tax rate is 10 percent for incomes below €10,000, and 30 percent for higher incomes. This
implies that $M_i(T^1)$ increases at a higher rate than $M_i(T^1)$ for $p_i > 0.36$. After the tax reform (tax system 2) the difference between the first and the second tax rate decreases (from 20 to 10 percent), so that the slope of the dashed red line $U_i(T^2)$ has no elbow around percentile 36, because $M_i(T^2)$ is higher than $M_i(T^1)$ for $p_i < 0.36$.

Figure 5: Zenga curves

Now we are interested in comparing the Kakwani effects measured according to the Lorenz and Zenga approaches. In Figure 6, the red lines show the differences between the ordinates of the Lorenz curves of gross income and tax liability for all $p_i$; analogously the green lines show the differences between the Zenga curve of gross income and tax liability in the two tax systems (solid line for tax system 1 and dashed line for tax system 2).

The red curves emphasize the cumulative effect due to the tax reform, and inform us about what is happening to the 91st percentile; the green ones tell us that the increase of the tax liability affects all taxpayers, starting from the poorest, and underline this with particular attention to the bottom part of the income distribution (the most affected by the tax reform), by contrasting it to the tax liability of the top income earners.
7. The data and the microsimulation model

To compare the Gini- and Zenga-based approaches when a real-world tax is considered, we employ the Bank of Italy *Survey on Household Income and Wealth* (hereafter SHIW) dataset, published in 2016. This survey contains information on household post-tax income and wealth in the year 2014, covering 8,156 households, and 19,366 individuals. The sample is representative of the Italian population, which is composed of about 24.7 million households and 60.8 million individuals. The SHIW provides information on each individual’s Personal Income Tax (PIT) net income, but does not contain the corresponding gross income; therefore, it has to be estimated for each taxpayer. For this purpose, we employ an updated version of the microsimulation model described in Morini and Pellegrino (2016).

A comparison of the results from the microsimulation model with the official statistics published by the Ministry of Finance (2016) shows that the distribution of gross income and of net tax according to bands of gross income and type of employment are in fact very similar. As a consequence, this microsimulation model is suitable for the empirical analysis we develop in the following section.
Finally, to perform our study, individual nominal incomes have to be transformed into equivalent incomes, using a proper equivalence scale. We choose to adopt the equivalence scale given by the square root of the number of the components of the household.

8. Results

We begin our analysis on the SHIW dataset by providing a description of the observed data. As a general picture, Figure 7 plots the Lorenz and Zenga curves for the pre-tax distribution as well as the concentration and Zenga curves for the post-tax distribution and the tax liability distribution. We see, for example, that the bottom 50 percent of households earns a mean gross income equal to roughly a quarter of the mean gross income of the top half.

Figure 7: Lorenz and Zenga U curves for gross and net incomes as well as tax distributions

![Figure 7](image)

Our purpose, here, is to analyze in greater detail the effect of a progressive taxation. Figure 8 plots the Lorenz and concentration curves for tax liability. For this purpose, the Lorenz-based approach is not so informative: the two curves lie approximately one above the other. By contrast, Figure 9 plots the same curves by employing the Zenga

15
approach: there is a greater relative difference between the two Zenga curves, when the ordering differs.

Figure 8: Lorenz and concentration curves for tax liability

By construction, the Zenga-based approach allows the observation of these phenomena, since the highest ordinate of the Zenga curve is lower than 0.15, in comparison to the
corresponding Gini-based approach, where the Lorenz and concentration curves always span from zero to 1.
Moreover, Figures 10 and 11 plot the $RS$ and $K$ effects, comparing the two approaches.

**Figure 10: $RS$ according to the Lorenz and Zenga approach**

The standard Lorenz-based approach, in both cases, generates curves that monotonically increase up to the 8\textsuperscript{th} decile, and then decrease. The Zenga approach tells us a different
story: the $RS$ effect, shown in Figure 10, increases up to the 4th decile, then it is more or less constant up to the 9th decile; differently, the $K$ curve shown in Figure 11 increases up to the 3th decile, it remains constant up to the 4th decile and then starts decreasing. Finally, similar plots for $REI_l$ can be easily obtained from the relationship involving the tax incidence for incomes greater than $x_i$; here we provided the ones for $RS$ and $K$, to show their value in interpreting data.

9. Concluding Remarks

In this paper we apply a recently proposed index of inequality, the Zenga index, to the study of the redistributive effect of taxation. Following the existing tax literature, we replicate the most important curves and the corresponding tax indexes (i.e., the Reynolds-Smolensky and the Kakwani index) by employing the new approach proposed by Zenga (2007). We also derive their mathematical relationship, in the new framework, along the lines of the earlier well-known equations, e.g., the Reynolds-Smolensky and the Kakwani indexes. We then underline the strengths and weaknesses of this approach when applied to the study of a tax reform.

The Zenga curve and index are strictly interconnected. The Zenga curve can be very useful for integrating information that cannot be inferred from the Lorenz curves. By contrasting the opposite parts of the distribution, the Zenga curve provides an insight into information that could be hidden in the cumulative approach underlined by the Lorenz curve. We have seen that the Kakwani and Reymolds Smolensky curves, redefined through the Zenga approach, seem to be more sensitive to an inequality reduction if it occurs in the bottom part of the distribution.

The research of this paper confirms that the Zenga approach is an innovative and very informative tool, particularly for highlighting what happens in different parts of the income distribution under the effects of a tax system.
References


