THE ZENGA EQUALITY CURVE: A NEW APPROACH TO MEASURING TAX REDISTRIBUTION AND PROGRESSIVITY

Published version:

DOI:10.1111/roiw.12493

(Article begins on next page)

The Zenga Equality Curve: A New Approach to Measuring Tax Redistribution and Progressivity

Francesca Greselin[*], Simone Pellegrino[†], Achille Vernizzi[‡]

October 2, 2020

Abstract

We adopt and extend the new Zenga inequality curve to study the degree of progressivity as well as the redistributive and re-ranking effects of a personal income tax system; moreover, we also establish the social welfare implications of these new inequality measures and compare them with the classical approach based on the Lorenz curve and the Gini coefficient. The Zenga methodology is based on comparing the mean income of the poorest income earners to the mean income of the remaining richest part of the population. To the best of our knowledge, this approach has never been applied to study the effects produced by a personal income tax. To fill this gap in the literature, we prove that the elasticity of the Zenga uniformity curve with respect to the Lorenz curve is always greater than 1, thus recasting – within the new paradigm – the most important curves and the corresponding tax indices, namely, the Reynolds-Smolensky, the Kakwani, and the Atkinson-Plotnick-Kakwani indices. We then derive three important inequalities for the newly developed measures, inspired by the well-known properties of the classical approach. Finally, we show how some information, which could remain unnoticed by the cumulative approach inherent to the Lorenz curve, is instead highlighted by the new methodology. The advantages of complementing the classic indices with the new ones are discussed through an application to the Italian tax system.

JEL-Codes: H23, H24.

Keywords : Personal Income Tax, Gini Index, bottom-to-top ratios, Zenga Index.

[*]Department of Statistics and Quantitative Methods, Università degli Studi di Milano-Bicocca. ORCID: 0000-0003-2929-1748. Phone: 0039 02 64 483 123. Email: francesca.greselin@unimib.it. Address: Via Bicocca degli Arcimboldi 8, 20126, Milano (IT).

[†]Corresponding Author. Department of Economics and Statistics – ESOMAS, Università degli Studi di Torino. ORCID: 0000-0001-8372-1054. Phone: 0039 011 670 6060. Email: simone.pellegrino@unito.it. Address: Corso Unione Sovietica 218bis, 10134, Torino (IT).

[‡]Department of Economics, Management, and Quantitative Methods, Università degli Studi di Milano. ORCID: 0000-0002-1641-5003. Phone: 0039 02 50 321 460. Email: achille.vernizzi@unimi.it. Address: Via Conservatorio 7, 20122, Milano (IT).

1. Introduction

Over more than a century of economic literature, several approaches have been proposed to study the inequality of quantitative variables, primarily income distributions.

Within this research stream, many synthetic indices appeared to summarise and compare the inequality of distributions by means of a single scalar. Among them, the most famous inequality index is, undoubtedly, the Gini (1914) coefficient which has also a graphical explanation through the Lorenz (1905) curve.

Additionally, a Gini-based methodology has been proposed in the tax literature to measure the degree of tax progressivity, yielding the (Gini-based) Kakwani (1977b) index, considering that relative income differentials get compressed in the transition from the pre- to the post-tax distribution. Similarly, the redistributive effect produced by the tax is measured by the (Gini-based) Reynolds and Smolensky (1977) index.

Although the Lorenz curve is a fundamental tool for welfare comparisons (Atkinson, 1970; Shorrocks, 1983; Atkinson and Bourguignon, 1987), we have found only a few applications of it in the tax literature, since it is not easy to draw specific conclusions by examining and comparing the Lorenz and concentration curves of different distributions (i.e. pre-tax distribution, post-tax distribution, and tax distribution). Indeed, due to their inherent cumulative nature, the different Lorenz curves are hardly distinguishable. Therefore, in most of the existing empirical research, the overall effect of taxation and transfer policy is primarily derived from the Gini and concentration coefficients (for a recent application see Guillaud et al. (2019)).

A few years ago, Zenga (2007) proposed a new methodology to plot and measure inequality, in which the new curve and index are based on comparisons between the mean income of the poorest income earners and the mean income of the remaining richest part of the population. Several recent studies have pointed out the different features of the Zenga approach with respect to the standard viewpoint, based on the Lorenz curve. Therefore, we move further in this literature stream, to explore the effectiveness of the former method in studying the effects of a personal income tax and the social welfare interpretations and implications of the tax progressivity (Son, 2013; Kakwani and Son, 2019) according to this new procedure. In particular, we mainly show that the graphical representation and some analytical tools based on the Zenga inequality curve provide an accurate instrument for understanding which part of a pre-tax distribution is mostly affected by the tax system or by a tax reform.

The remainder of this paper is organized as follows. Section 2 briefly reviews the Lorenz and Gini approaches and then recalls their implications about social welfare. Similarly, Section 3 outlines the new Zenga curve and index, reminds their properties, and presents the social welfare implications within the Zenga approach. In view of our main purpose, that is, to analyse the effects of taxation on income distributions, Section 4 extends the Zenga curve and index, defining three new tools for measuring the degree of progressivity of a personal income tax, as well as its re-ranking and redistributive effects. Important relationships among the new indices and curves have been derived, and the social welfare implications of taxation are discussed. Section 5 shows an application to a tax case; in particular, Subsection 5.1 introduces the microsimulation model employed for the empirical estimations, whilst Subsection 5.2 develops a compared analysis with the new and standard approach, to show the interest and potential of the proposal; finally, Subsection 5.3 presents the social welfare implications. Section 6 offers some concluding remarks. We gather in Appendix A some technical proofs and in Appendix B formulas for the derivations of empirical curves and indices using survey data.

2. The Lorenz curve and the Gini index

2.1. Basic notation

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 the corresponding population quantile function by $F^{-1}(p) =: z_p = inf\{z : F(Z) \geq p\}$, for $0 < p < 1$.

The Lorenz (1905) curve plots 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 $(1, 1)$. Conversely, 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 a zero income), the share of incomes equals zero for $0 \leq p < 1$, so that $L_F(p) = 0$, and only for $p = 1$ we have $L_F(1) = 1$. It is given by $(p, L_F(p))$, where

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

and $\mu_F = E(Z)$ 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 line we get in the case of perfect equality, namely, the diagonal line. The Gini coefficient is precisely given by twice the area between the equality line and the Lorenz curve[1]

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

Now, considering the mean income of the poorer $p\%$ of the population

$$\mu_F^-(p) = \frac{1}{p}\int_0^p F^{-1}(s)ds, \tag{3}$$

the Gini index can be rewritten in terms of the relative deviation of the mean income of the poorer $p\%$ of the population from the overall mean μ_F:

$$G_F = \int_0^1 \frac{\mu_F - \mu_F^-(p)}{\mu_F}\,2p\,dp. \tag{4}$$

2.2. Social welfare implications

Following the seminal paper by Atkinson (1970) and related literature (see Son (2013) and Kakwani and Son (2019) for a review of the literature and extensions), government policies should be judged based on their impact on social welfare, which is an aggregate measure of society's welfare derived from each individual's welfare levels. Within this framework, the area under twice the (generalized)

[1] In the original works (Gini, 1912, 1914; Pietra, 1915), the coefficient is given in terms of discrete distributions.

Lorenz curve can be interpreted as follows

$$W_G = 2\int_0^1 \mu L(p)dp = \mu(1-G) = 2\int_0^{+\infty} z_p(1-p)dp. \quad (5)$$

Eq. (5) is the social welfare function implied by the Gini index that was proposed by Sen (1974). Note from the third term on the right hand side of Eq. (5) that the Gini social welfare function is the weighted average of individual welfare levels z_p, with weight given by $w(p) = 2(1-p)$. It can easily be shown that the total weight adds up to 1 for the entire population. Furthermore, the weight is proportional to the welfare ranking of individuals: the poorest individual receives the maximum weight and the richest individual gets the minimum weight. If this pro-poor weighting of welfare is acceptable to policymakers, then it is a useful tool to analyze government policies (see Kakwani and Son (2019)).

3. The new Zenga curve and index

3.1. Basic notation

The increasing gap observed between the less fortunate and the more fortunate individuals (see, among many others, Piketty (2013)), motivated a fresh rethinking with respect to inequality and gave rise to many proposals in the literature (see Zenga (2007); Gastwirth (2014, 2016); Davydov and Greselin (2018), and references wherein). There is a consensus that no measure can be considered superior to the others (Osberg (2017)); therefore, the choice of an inequality measure must rest on its appropriateness for specific substantive problems (Jasso, 1982). In particular, a few years ago, to capture the recent changes in the extreme parts of the income distribution, Zenga (2007) proposed a new inequality curve $I_F(p)$ based on contrasting the average income of the poorer $p\%$ bottom earners $\mu_F^-(p)$ defined in (3) with the amount that is held, on average, by the richest top earners, namely, the remaining $(1-p)\%$ of the population:

$$\mu_F^+(p) = \frac{1}{1-p}\int_p^1 F^{-1}(s)ds.$$

Therefore, Zenga (2007) defined the curve $(p, I_F(p))$ where

$$I_F(p) = \frac{\mu_F^+(p) - \mu_F^-(p)}{\mu_F^+(p)}$$

for $0 < p < 1$. When the random variable Z is equal to a constant, the corresponding quantile $F^{-1}(p)$ is also equal to the constant, as well as both the lower and the upper means $\mu_F^-(p)$ and $\mu_F^+(p)$; thus, $I_F(p) = 0\ \forall p \in (0,1)$, indicating perfect equality or an 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; thus, $I_F(p) = 1\ \forall p \in (0,1)$. As illustrated by Greselin et al. (2010), this approach "*considers that the notions of poor and rich are relative to each other*" and summarises, in a single measure, the amount of inequality in the population by the following index:[2]

$$I_F = \int_0^1 \frac{\mu_F^+(p) - \mu_F^-(p)}{\mu_F^+(p)}\, dp. \tag{6}$$

3.2. Properties of the Zenga index

The Zenga index obeys a number of properties which can be regarded as being intrinsic to the concept of inequality. In what follows, we review such properties, using the notation I_Z or I_F for the index, whenever this helps in simplifying the presentation.

i) *Scale Invariance.* The Zenga inequality index is a relative inequality measure, as proportional changes in all incomes, where changing Z in cZ with $c > 0$, does not alter the level of inequality:

$$I_Z = I_{cZ}. \tag{7}$$

Technically, we say that it is homogeneous of degree zero in incomes. Evidently, an index satisfying (7) handles the money illusion , namely, if incomes are measured in pounds instead of dollars, then inequality does not change.

ii) *Sensitivity to translation.* By adding any constant $c > 0$ to the income Z, transforming Z into $Z + c$, the relative measure of inequality decreases:

$$I_{Z+c} \leq I_Z.$$

In other words, a relative measure of inequality has to consider that the relative distances within incomes are reduced by adding a constant positive amount.

iii) *Lorenz ordering.* Following Aaberge (2000), the Lorenz ordering $Z \leq_L Y$ indicates the bound $L_Z(p) \geq L_Y(p)$ for all $p \in [0, 1]$. If the random variables Z and Y follow the Lorenz ordering, then

$$I_Z \geq I_Y.$$

iv) *Pigou-Dalton transfer principle.* Pigou-Dalton transfer principle states that progressive (i.e., from rich to poor) rank-order and mean-preserving transfers should decrease the value of inequality measures, changing Z into Y, yielding

$$I_Z \geq I_Y.$$

Comparing the Lorenz and the Zenga curves,[3] we observe that while the Zenga $I_F(p)$ contrasts

[2] At the time of paper submission, we found a similar approach in Jasso (2018), based on the ratio of average income among the top 1% to the average income among the bottom 99%, called "TopBot". The author states that "*TopBot is not constrained to the 99-1% split, but can be used with any percentage split of a population into two sub-populations, such as 90-10, 50-50, 10-90, 1-99, and so on, or whatever may be most appropriate for the substantive context.*" To complete the comparison with Jasso (2018), the Zenga (2007) index integrates the standard form of TopBot as: "one minus the bottom-to-top ratio" $1 - \mu^-(p)/\mu^+(1-p)$ over all percentiles p for the Bottom group, and the complementary group of $1 - p$ for the Top group.

[3] Even though the literature on the Zenga index and curve is obviously not as extensive as the one on the Gini index, we find research on various features of the index and curve (Polisicchio, 2008; Polisicchio and Porro, 2009; Jedrzejczak and Trzcinska, 2019; Maffenini and Polisicchio, 2014; Greselin et al., 2009; Arcagni and Porro, 2014), inferential results and their applications (Greselin and Pasquazzi, 2009; Greselin et al., 2010, 2013, 2014), subgroup decompositions of the index (Radaelli, 2008, 2010), longitudinal decomposition (Mussini and Zenga, 2013) decompositions by income sources (Zenga et al., 2012; Pasquazzi and Zenga, 2018), and many applications on real data (Arcagni and Zenga, 2013). The interested reader can find accurate discussions of the advantages of the Zenga index over the Gini index within both the descriptive and inferential frameworks. Langel and Tillé (2012) analyzed the sampling distribution of the empirical Zenga index, and Antal et al. (2011) extended inferential results to complex sampling designs.

the mean incomes of two disjoint and exhaustive sub-populations (i.e. the poor and the rich), $L_F(p)$ compares the economic condition of the poorer group to that of the entire population.

Finally, the most important consideration is related to the weight function $2p$ adopted to define the Gini index in (4), for normalization purposes. It places a lower emphasis on the most crucial comparisons referring to the poorest portions of populations, and a greater emphasis on the comparisons between almost coinciding sub-populations, which are likely to be less informative[4]. However, the Zenga index considers, with the same weight, any relative deviation from equality, measured by $I_F(p)$, in any part of the distribution.

3.3. A comparison between the Zenga and the Lorenz curves

In the analysis of real data and the effects of different taxation systems, we prefer to switch from the Zenga inequality curve to its complement to 1, that is, the Zenga uniformity curve. In the following Sections, the latter plays a central role. It is given by $(p, U_F(p)) = (p, 1 - I_F(p))$, where

$$U_F(p) = \frac{\mu_F^-(p)}{\mu_F^+(p)}.$$

We introduced $I(p)$ as an inequality curve, being equivalent to the claim that $U(p)$ is a curve measuring the extent of equality across the opposite groups of poorer and richer individuals in the population.[5] Any departure from the perfect equality is represented by the deviation of the uniformity curve from 1. Moreover, $U(p)$ has a more direct interpretation, as we will see in the application to real economic data in Section 5. The Lorenz curve itself can be expressed in terms of means and percentiles as

$$L_F(p) = \frac{p\ \mu_F^-(p)}{\mu_F},$$

which also yields

$$L_F(p) = U_F(p)\ \frac{p\ \mu_F^+(p)}{\mu_F}. \tag{8}$$

To analyse the sensitivity of the uniformity curve with respect to a variation in the Lorenz curve we start from another equivalent formulation of their analytical relationship (Zenga, 2007)

$$U_F(p) = \frac{(1-p)L_F(p)}{p(1-L_F(p))},$$

and we obtain

$$\frac{\partial U_F(p)}{\partial L_F(p)} = \frac{(1-p)}{p(1-L_F(p))^2}.$$

The elasticity measures the proportional change in an economic variable in response to a change in another; therefore, the elasticity of the uniformity curve with respect to the Lorenz curve is given

[4]Due to this consideration, many generalization of the Gini index arose in the literature, emphasising or de-emphasising, depending on the problem under consideration, the difference $p - L(p)$ in some regions of the unit interval [0, 1]; among them, we cite Donaldson and Weymark (1980) and Yitzhaki (1983).

[5]An anonymous referee suggested us to employ the word "equality" to designate the curve $U(p)$, being "equality" the opposite of "inequality". We opted for the name given to it by Zenga, in the original paper: "$U(p)$ measures the *uniformity* between the lower and the upper group" of poorer and richer, respectively.

by

$$\frac{\partial U_F(p)/\partial L_F(p)}{U_F(p)/L_F(p)} = \frac{1}{1-L_F(p)} > 1. \tag{9}$$

This means that an increase of 1% in the Lorenz ordinate causes a change in the ordinate of the uniformity curve greater than 1%, for all $p \in (0,1)$. Moreover, we see that the elasticity in (9) increases as $L_F(p)$ approaches 1, that is, with the increase in percentile p. An increase of 1% in the income of the poorest p% can be hardly noticed as a difference between the two Lorenz curves. However, due to the greater eleasticity of $U_F(p)$ with respect to $L_F(p)$, the same 1% increase in the income of the poorest p% may produce a visible shift in $U_F(p)$. We will empirically observe the effects of this higher sensitivity of the Zenga uniformity curve in Section 5, where the latter reacts with more evidence to a new tax system, generating a modification in the underlying distribution F.

The Lorenz curve captures, in a sense, the essence of inequality, by displaying the deviation of each person's welfare from perfect equality. The nearer the Lorenz curve is to the egalitarian line, the more equal the distribution of welfare. We end this section by showing the equivalence between Lorenz dominance, and the ordering based on the $U_F(p)$ curve. Let F_X and F_Y be the distribution functions of the continuous non-negative random variables X and Y, both with finite and positive expected value. We need here to introduce some notions of dominance:

Definition 1. *We say that X dominates Y under the Lorenz ordering, denoting it by $F_X \geq_L F_Y$ if and only if $L_{F_X}(p) \leq L_{F_Y}(p) \quad \forall p \in (0,1)$.*

Definition 2. *We say that X dominates Y under the Uniformity ordering, denoting it by $F_X \geq_Z F_Y$ if and only if $U_{F_X}(p) \leq U_{F_Y}(p) \quad \forall p \in (0,1)$.*

Proposition 3. *The Lorenz ordering and the ordering based on the Uniformity curve are equivalent.*[6]

Proof. We may rewrite (8) as follows

$$L_F(p) = 1 - \frac{1-p}{1-p+p\,U_F(p)}.$$

Therefore, $L_{F_X}(p) \leq L_{F_Y}(p) \iff L_{F_X}(p) = 1 - \frac{1-p}{1-p+p\,U_{F_X}(p)} \leq 1 - \frac{1-p}{1-p+p\,U_{F_Y}(p)}, \iff U_{F_X}(p) \leq U_{F_Y}(p)\ \forall p \in (0,1)$. □

Consequently, the Uniformity curve can be used as a criterion for ranking government policies or programs.

3.4. Social welfare implications

Along the lines of Son (2013) and Kakwani and Son (2019), we introduce here the social welfare evaluation based on the Zenga Uniformity curve

$$\begin{aligned} W_Z &= \mu \int_0^1 \frac{L_F(p)}{(1-L_F(p))} \frac{(1-p)}{p}\, dp = \mu(1-Z) \\ &= \int_0^1 z_p\, \frac{(-\ln p + p - 1)}{(1-L_F(p))^2}\, dp \end{aligned} \tag{10}$$

[6]Polisicchio and Porro (2009) have shown that the ordering based on the Zenga inequality curve is equivalent to the Lorenz ordering.

where the weight function $w_Z(p)$ is obtained after integrating by parts W_Z, as follows

$$w_Z(p) \;=\; \frac{(-\ln p + p - 1)}{(1 - L_F(p))^2}.$$

We can verify that the function described by Eq. (10) is homogeneous of degree one, implying that if we change all incomes by the same proportion, this function also varies by the same proportion. We then decompose $w_Z(p)$ into two multiplicative terms

$$\begin{aligned} w_Z(p) \;&=\; \frac{(-\ln p + p - 1)}{(1-p)^2}\,\frac{(1-p)^2}{(1 - L_F(p))^2} \\ &:=\; w_Z^*(p)\;\beta_Z(p). \end{aligned}$$

The first term

$$w_Z^*(p) = \frac{(-\ln p + p - 1)}{(1-p)^2},$$

is a *non negative, concave upward* and *strictly decreasing* function of the rank p, like the social welfare functions

$$w_G(p) = [2\,(1-p)], \qquad w_{G_k}(p) = [k\,(1-p)^k], \qquad \text{and} \qquad w_B(p) = [-\ln p]$$

that are implicit in the Gini index G, the generalized Gini G_k, and the Bonferroni index B, respectively (see Kakwani and Son (2019)). It can easily be shown that the total weight $w_Z^*(p)$ adds up to 1. Moreover, if everyone receives the same income, then the social welfare function W_Z must be equal to μ. To satisfy this requirement, the total weight implied by $w_Z(p)$ must add up to one. This is verified since, when all incomes are equal, we have that $L_F(p) = p$, so that

$$w(p) = \frac{-\ln p + p - 1}{(1 - L_F(p))^2} = \frac{-\ln p + p - 1}{(1-p)^2} = w^*(p). \tag{11}$$

The function $w_Z^*(p)$ incorporates a society's distributional judgement, where the poorest individual receives the maximum weight and the richest individual gets the minimum weight (see Greselin et al. (2020)). The second term

$$\beta_Z(p) = \frac{(1-p)^2}{(1 - L_F(p))^2} = \left(\frac{\mu}{\mu_F^+(p)}\right)^2$$

depends, instead, on both p and $L(p)$. $\beta_Z(p)$ is a decreasing function of p, with $\beta_Z(0) = +1$ and $\lim_{p\to 1} \beta_Z(p) = 0$. The greater the ratio $\frac{\mu_F^+(p)}{\mu}$ the greater the penalization given to the income z_p by $\beta_Z(p)$. Comparing to what we have recalled about the Gini and Bonferroni indexes, in the social welfare evaluation based on the Zenga Uniformity curve, beyond to the weight function based on the ranks p, we also have $\beta_Z(p)$. This difference arises from the fact that the denominator for the Gini and the Bonferroni indexes is a constant value, while in the present approach the denominator is a function of p.

4. The Lorenz and Zenga approach, considering tax effects

4.1. General overview

In the following, we jointly analyse the pre- and post-tax income, as well as the tax distribution. First, we briefly review the standard approach based on the Lorenz curves and the Gini coefficients, and then we construct analogous curves and indices by extending the Zenga approach.

We assume that the pre-tax incomes $\tilde{x} = \{x_1, x_2, \ldots, x_n\}$ are arranged in non-decreasing order (i.e. $x_i < x_{i+1}$ for $i = 1, \ldots, n-1$). Let $T(x)$ be the tax paid by an individual of income x. The post-tax or disposable income of the individual will then be $y(x) = x - T(x)$. Let us consider the after-tax incomes $\tilde{y} = \{y_1, y_2, \ldots, y_n\}$ and taxes $\tilde{t} = \{t_1, t_2, \ldots, t_n\}$, so that each triplet (x_i, y_i, t_i) refers to the $i - th$ individual of the sample. Since, for each pair of individuals i, j such that $x_i < x_j$, it is not granted that $y_i < y_j$ and $t_i < t_j$, we denote the post-tax incomes by Y when they are ordered in a non-decreasing order, and the same incomes by Y_X when units are ranked according to the pre-tax order. Similarly, we will denote the observations related to tax amounts by T and T_X.

Based on the seminal work by Musgrave and Thin (1948), several indices have been constructed to evaluate the redistributive and re-ranking effects and the degree of progressivity of a tax system. They are mainly functions of the Gini coefficients $G(X)$, $G(Y)$, and $G(T)$ and of the corresponding concentration (or pseudo-Gini) coefficients $C(Y_X)$ and $C(T_X)$.

Therefore, it may be useful to recall the definition of the concentration curve and coefficient (Kakwani, 1977a). In Section 2, the Lorenz curve has been defined as the relationship between the proportion p of the population having income less than or equal to x, and the corresponding proportion $L_F(p)$ of owned income, $(p, L_F(p))$. Here, we need to recast the Lorenz curve in terms of the income x, by exploiting the equality $p = F(x)$. Thus, $L_F(p) = L_F(F^{-1}(x))$ is a function of x with values in $[0, 1]$, as a genuine *cdf*, which we may denote by $F_1(x)$ for short, and therefore, the Lorenz curve can be equivalently expressed by $(F(x), F_1(x))$.

Let us now consider the variable W, observed on the same sample (or population), and let $F_1[W(x)]$ be the share of W owned by the statistical units having a value of X less than or equal to x (i.e. $F_1[W(x)]$ cumulates the values of W along the ordering given by X). The concentration curve of W is the relationship between $F(x)$ and $F_1[W(x)]$, and the concentration index $C(W)$ is equal to one minus twice the area under the concentration curve.

Now, we are ready to recall the most important indices for analysing tax effects.

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 (Jakobsson, 1976). If the tax elasticity is equal to 1 at all income levels x, the two cdf curves $F_1(x)$ and $F_1(T(x))$ coincide, as the greater the distance between them, the larger the difference of the tax elasticity from unity.

The overall degree of progressivity is generally evaluated by the Kakwani (1977b) index

$$K = C(T_X) - G(X),$$

that is twice the area between the Lorenz curve of X and the concentration curve of T_X. Therefore, K measures the departure from proportionality of the actual tax system.

Similarly, we introduce an analogous tool for measuring tax progressivity, in accordance with the Zenga approach. To this end, we need to develop the concepts of the concentration curve and index within the Zenga approach. We evaluate the concentration of the taxes, and denote the generic ordinate of the tax uniformity curve by $U_{T_X}(p)$ and the corresponding index by $U(T_X)$, considering the taxes amounts $\{t_i\}$, not in their natural sequence, but when the latter are sorted by the ordering induced by the sorted incomes $\{x_i\}$.[7] Now, we are ready to introduce the new curve $KI(p)$ and the synthetic measure KI, derived by the Zenga approach:

$$KI(p) = I_{T_X}(p) - I_X(p) = U_X(p) - U_{T_X}(p)$$

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

We see that $KI(p)$ involves differences between the ordinates of the tax uniformity curve $U_{T_X}(p)$ and the pre-tax uniformity Zenga curve $U_X(p)$. 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. The difference between the Gini coefficients of the pre-tax X and post-tax Y income distributions assesses the overall redistributive effect, measured by the RE index:

$$RE = G(X) - G(Y) = \big(G(X) - C(Y_X)\big) - \big(G(Y) - C(Y_X)\big).$$

If we compare the Gini coefficient of the pre-tax distribution X, and post-tax concentration of Y_X, considering the units sorted according to the pre-tax incomes in both cases, we arrive at the Reynolds-Smolensky index RS:

$$RS = G(X) - C(Y_X). \tag{12}$$

Hence, the overall 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 is given by twice the area between the Lorenz curve for the pre-tax distribution and the concentration curve for the post-tax distribution (Reynolds and Smolensky, 1977; Lambert, 2001).

If the tax determines re-ranking, then $G(Y) > C(Y_X)$ and $RS > RE$. The Atkinson-Plotnick-Kakwani index (Atkinson, 1980; Plotnick, 1981; Kakwani, 1984) is a measure of the overall re-ranking in the transition from pre- to post-tax income distribution, defined by:

$$R(Y_X) = G(Y) - C(Y_X). \tag{13}$$

Similarly, as we did before with taxes, beyond the uniformity curve and index $U_Y(p)$ and $U(Y)$ evaluated on the sorted values $\{y_i\}$, the concentration curve $U_{Y_X}(p)$ and index $U(Y_X)$ arise when we consider the post-taxes amounts $\{y_i\}$, not in their natural sequence, but when the latter are sorted by the ordering induced by sorted incomes $\{x_i\}$.

Following the Zenga approach, we introduce three new curves and synthetic indices:

$$REI(p) = I_X(p) - I_Y(p) = U_Y(p) - U_X(p) = RSI(p) - RI(p)$$

$$REI = I(X) - I(Y) = U(Y) - U(X) = RSI - RI$$

$$RSI(p) = I_X(p) - I_{Y_X}(p) = U_{Y_X}(p) - U_X(p)$$

[7] It holds that $U_{T_X}(p) \geq U_T(p)$ and, equivalently, $I_{T_X}(p) \leq I_T(p)$ because $\mu^-_{T_X} \geq \mu^-_T$ and $\mu^+_{T_X} \leq \mu^+_T$.

$$RSI = I(X) - I(Y_X) = U(Y_X) - U(X)$$

$$RI(p) = I_Y(p) - I_{Y_X}(p) = U_{Y_X}(p) - U_Y(p)$$

$$RI = I(Y) - I(Y_X) = U(Y_X) - U(Y). \tag{14}$$

Three important and well-known results hold for the classical measures of the degree of progressivity, the redistributive and the re-ranking effects of a tax system. First, we recall them, and then derive their analogous counterparts in the new setting.

i) The Kakwani progressivity index K is related to the Reynolds-Smolensky RS index by (see, e.g. Lambert (2001)):

$$RS = K\ \overline{T}/\overline{Y},$$

where $\overline{T}$ and $\overline{Y}$ are the averages of T and Y, respectively.[8] This means that RS is a function of two variables (i.e. 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 for the effect of the tax rate. Thus, an analogous relation holds between RSI and KI, based on the corresponding curves (see the Appendix)

$$RSI(p) = I_X(p) - I_{Y_X}(p) = \lambda(p)\ \left[I_{T_X}(p) - I_X(p)\right] = \lambda(p)\ KI(p). \tag{15}$$

In (15), $KI(p) = I_{T_X}(p) - I_X(p)$ measures the tax progressivity at percentile p, while the factor $\lambda(p) = \mu^+_{T_X}\ \mu^+_{Y_X}$ measures the tax incidence on incomes greater than x_p; note that the term $\lambda(p)$, which multiplies $\left(I_{T_X}(p) - I_X(p)\right)$, is generally different for each p: if the tax is progressive, $\lambda(p)$ is an increasing function of X, reflecting the tax system progressivity. In analogy with Eq. (12), the synthetic measure based on the curve in (15) can be expressed as $RSI = I(X) - I(Y_X)$.

ii) The Atkinson-Plotnick-Kakwani takes only non-negative values. In analogy with this property, we will show that $RI \geq 0$. By recalling that $\mu^-_{Y_X}(p) \geq \mu^-_Y(p)$ and $\mu^+_{Y_X}(p) \leq \mu^+_Y(p)$, since Y is in non-decreasing order, then

$$\begin{aligned} RI(p) &= \frac{\mu^+_Y(p) - \mu^-_Y(p)}{\mu^+_Y(p)} - \frac{\mu^+_{Y_X}(p) - \mu^-_{Y_X}(p)}{\mu^+_{Y_X}(p)} \\ &= \frac{\mu^-_{Y_X}(p)}{\mu^+_{Y_X}(p)} - \frac{\mu^-_Y(p)}{\mu^+_Y(p)} \geq 0. \end{aligned}$$

The last inequality strictly holds if and only if the tax system induces some re-ranking effect among the poorest p percent individuals in the population; otherwise, $RI(p) = 0$. This property gives RI the role of a measure of the re-ranking effect produced by the tax.

iii) In analogy with the well-known inequality $RS \geq RE$, we have that $RSI(p) \geq REI(p)$ for all

[8] As we will discuss later (see Subsection 4.2), $\overline{T}/\overline{X} = \alpha$ and $\overline{T}/\overline{Y} = \alpha/(1-\alpha)$.

p, and therefore:

$$RSI \geq REI. \tag{16}$$

To prove our claim, by conveniently decomposing $REI(p)$ as follows

$$\begin{aligned} REI(p) &= I_X(p) - I_Y(p) \\ &= [I_X(p) - I_{Y_X}(p)] - [I_Y(p) - I_{Y_X}(p)] \\ &= RSI(p) - RI(p), \end{aligned}$$

we get the thesis using ii). Hence, the equality in (16) holds if and only if the tax does not determine the re-ranking in Y.

4.2. Comparing the Lorenz- and Zenga-based approaches for tax analysis

To understand the different behaviours of the curves $RS(p)$ and $RSI(p)$ as well as $K(p)$ and $KI(p)$, and to emphasise their relationships, we get a step further by decomposing the corresponding equations. Similar to $\mu_Z^-(p)$ and $\mu_Z^+(p)$, we define the lower and upper average tax rates

$$\alpha^-(p) = \frac{\mu_T^-(p)}{\mu_X^-(p)} \tag{17}$$

$$\alpha^+(p) = \frac{\mu_T^+(p)}{\mu_X^+(p)}. \tag{18}$$

Additionally, we denote the overall pre-tax mean income by $\overline{X}$, and the overall average tax rate by $\alpha = \overline{T}/\overline{X}$.

Thus, we can re-interpret $RS(p)$, $RSI(p)$, $K(p)$ and $KI(p)$ as follows:

$$RS(p) = \frac{\mu_X^-(p)}{\overline{X}} \, \frac{\alpha - \alpha^-(p)}{1-\alpha} \, p = L_X(p) \frac{\alpha - \alpha^-(p)}{1-\alpha} \tag{19}$$

$$K(p) = \frac{\mu_X^-(p)}{\overline{X}} \, \frac{\alpha - \alpha^-(p)}{\alpha} \, p = L_X(p) \frac{\alpha - \alpha^-(p)}{\alpha} \tag{20}$$

$$RSI(p) = \frac{\mu_X^-(p)}{\mu_X^+(p)} \, \frac{\alpha^+(p) - \alpha^-(p)}{1-\alpha^+(p)} = U_X(p) \frac{\alpha^+(p) - \alpha^-(p)}{1-\alpha^+(p)} \tag{21}$$

$$KI(p) = \frac{\mu_X^-(p)}{\mu_X^+(p)} \, \frac{\alpha^+(p) - \alpha^-(p)}{\alpha^+(p)} = U_X(p) \frac{\alpha^+(p) - \alpha^-(p)}{\alpha^+(p)}. \tag{22}$$

In the Lorenz-Gini approach, the redistributive effect $RS(p)$ in (19) and the degree of progressivity $K(p)$ in (20) depend, for each p, on three elements: the ordinates of the Lorenz curve for the pre-tax income distribution $L_i(X)$, the overall average tax rate α, and the average tax rate obtained by considering only the bottom $p\%$ income units $\alpha^-(p)$.
Moreover, we notice that the ratio between $RS(p)$ and $K(p)$ is constant, since $\frac{RS(p)}{K(p)} = \frac{\alpha}{1-\alpha}$. Therefore, the two curves convey the same information; however, the same cannot be observed for the ratio between $RSI(p)$ and $KI(p)$, since

$$\frac{RSI(p)}{KI(p)} = \frac{\alpha^+(p)}{1-\alpha^+(p)}. \tag{23}$$

The Zenga approach offers a different tool, since in (21) and (22) there is no constant term; thus, both $RSI(p)$ and $KI(p)$ contribute to explaining the impact of the tax, with each of them conveying distinct information. The redistributive effect $RSI(p)$ and the degree of progressivity $KI(p)$ depend, for each p, on three elements: the ordinates of the Zenga curve for the pre-tax income distribution $U_X(p)$, the average tax rate for the bottom $p\%$ income units $\alpha^-(p)$, and the average tax rate for the top $(1-p)\%$ income units $\alpha^+(p)$.

We have seen that the Lorenz curve in (19) and (20), is replaced by the uniformity curve in (21) and (22). Additionally, the uniformity curve lies over the Lorenz curve for the lowest percentiles, whereas the Lorenz curve lies over the uniformity curve for the highest percentiles. Evidently, there exist two values $p^*, p^{**} \in (0,1)$ such that $U_X(p) \geq L_X(p)$ for all $p < p^*$ and $U_X(p) \leq L_X(p)$ for $p > p^{**}$ (the proof is given in the appendix).

Morevover, in the Gini-Lorenz approach in (19) and (20), $\alpha^-(p)$ is compared with the overall tax rate α, while in the Zenga-based approach in (21) and (22), the comparison is made the average tax rate for the higher percentiles $\alpha^+(p)$.

Furthermore, $\frac{\alpha-\alpha^-(p)}{1-\alpha}$ and $\frac{\alpha-\alpha^-(p)}{\alpha}$ decrease with p, while the ratios $\frac{\alpha^+(p)-\alpha^-(p)}{1-\alpha^+(p)}$ and $\frac{\alpha^+(p)-\alpha^-(p)}{\alpha^+(p)}$ generally have a more complex behaviour. In terms of their interpretation, $\frac{\alpha-\alpha^-(p)}{\alpha}$ represents the (relative) comparison between the overall average tax rate α and the average tax rate of the poorest $p\%$ income earners. Conversely, the ratio $\frac{\alpha^+(p)-\alpha^-(p)}{\alpha^+(p)}$ represents the (relative) comparison between the average tax rate of the poorest $p\%$ income earners and the average tax rate of the upper $(1-p)\%$ part of the distribution (i.e. the richest income earners over the percentile p).

We may add a second interpretation of the same two ratios, in terms of disposable incomes. Since $\alpha - \alpha^-(p) = (1-\alpha^-(p)) - (1-\alpha)$ and $\alpha^+(p) - \alpha^-(p) = (1-\alpha^-(p)) - (1-\alpha^+(p))$, $\frac{\alpha-\alpha^-(p)}{1-\alpha}$ is the difference between the average disposable income after taxation, up to percentile p and the average disposable income of the entire population. On the contrary, in $\frac{\alpha^+(p)-\alpha^-(p)}{1-\alpha^+(p)}$, the average disposable income after taxation, up to percentile p, is compared to the average disposable income of the complementary $(1-p)\%$ richest part of the population. In other words, we propose to assess the redistribution and progressivity of a tax system by comparing their effects on the poorest and richest segments of the population.

4.3. Social welfare implications of taxation

Following the detailed analysis shown in Kakwani and Son (2019), in this section we present how social welfare gains and losses due to taxation can be measured and quantified. Basically, the cited Authors propose "*a social welfare function framework to derive measures of tax progressivity and explore their normative properties.*" These measures can be absolute and relative: absolute measures evaluate the deviation of an observed tax system to a situation in which all tax units pay the same amount of taxes; relative ones measure the deviation of an observed tax system to a situation in which all tax units are affected by the same average tax rate. Moreover, by considering increasing and concave social welfare functions, and functions characterized by homogeneity of degree one, evaluated social welfare levels have a money level interpretation.

In particular, focusing on Eq. (5) and (10), we denote by W_G^X and W_Z^X the overall social welfare levels when pre-tax incomes are considered, where the labels G and Z refer to the Gini and the Zenga approach. Similarly, we call W_G^Y and W_Z^Y the overall social welfare levels when post-tax incomes are considered according to the post-tax order and $W_G^{Y|X}$ and $W_Z^{Y|X}$ the overall social welfare levels when post-tax incomes are considered according to the pre-tax order.

Then Kakwani and Son (2019) derive specific equations for the social welfare gains and losses of taxation and their decompositions. In particular, $W_G^Y - W_G^X < 0$ and $W_Z^Y - W_Z^X < 0$ measure the absolute difference in social welfare levels after and before taxation, whilst the corresponding relative measures, $\frac{W_G^Y - W_G^X}{\bar{T}} < 0$ and $\frac{W_Z^Y - W_Z^X}{\bar{T}} < 0$, are simply obtained by dividing the previous values by the average tax amount $\bar{T}$.

Three relevant relations holds for the Gini approach: $W_G^X = \bar{X}(1-G(X))$; $W_G^Y = \bar{Y}(1-G(Y))$; $W_G^{Y|X} = \bar{Y}(1 - C(Y_X))$. A same set of equations can be derived also for the Zenga methodology: $W_Z^X = \bar{X}(1 - I(X))$; $W_Z^Y = \bar{Y}(1 - Z(Y))$; $W_Z^{Y|X} = \bar{Y}(1 - I(Y_X))$.

Having defined the overall average tax rate by $e = \frac{\bar{T}}{\bar{X}}$, Kakwani and Son (2019) decompose the absolute loss of welfare due to taxation, $W_G^Y - W_G^X$, in three parts, as follows:

$$W_G^Y - W_G^X = H_G^A - eW_G^X + \left(W_G^{Y|X} - (1-e)W_G^X\right). \tag{24}$$

For the Zenga approach we similarly obtain

$$W_Z^Y - W_Z^X = H_Z^A - eW_Z^X + \left(W_Z^{Y|X} - (1-e)W_Z^X\right). \tag{25}$$

In particular,

i) $H_G^A = W_G^Y - W_G^{Y|X} = -\bar{Y}R(Y_X) < 0$ and $H_Z^A = W_Z^Y - W_Z^{Y|X} = -\bar{Y}RI < 0$ capture the social welfare loss due to re-ranking: therefore, they are measures of absolute horizontal inequity;

ii) $eW_G^X = \bar{T}(1 - G(X)) > 0$ and $eW_Z^X = \bar{T}(1 - I(X)) > 0$ can be thought as the social welfare loss when adopting a proportional tax that yields the same tax revenue of the actual progressive tax, and

iii) $W_G^{Y|X} - (1-e)W_G^X = \bar{Y}RS > 0$ and $W_Z^{Y|X} - (1-e)W_Z^X = \bar{Y}RSI > 0$ measure the progressivity of the tax system.

The corresponding relative decompositions and measures are simply obtained by dividing each term by the average tax amount $\bar{T}$ as earlier mentioned. As a consequence, decomposing $\frac{W_G^Y - W_G^X}{\bar{T}}$ and $\frac{W_Z^Y - W_Z^X}{\bar{T}}$, we obtain

$$H_G^R = \frac{W_G^Y - W_G^{Y|X}}{\bar{T}} = -\frac{\bar{Y}}{\bar{T}}R(Y_X) \qquad \text{and} \qquad H_Z^R = \frac{W_Z^Y - W_Z^{Y|X}}{\bar{T}} = -\frac{\bar{Y}}{\bar{T}}RI,$$

they basically measure the average loss of social welfare due to reranking for one euro of tax collected by the Government. Similarly, we have

$$\frac{eW_G^X}{\bar{T}} = \frac{W_G^X}{\bar{X}} = 1 - G(X) \qquad \text{and} \qquad \frac{eW_Z^X}{\bar{T}} = \frac{W_Z^X}{\bar{X}} = 1 - I(X),$$

and we arrive, finally, at

$$\frac{W_G^{Y|X} - (1-e)W_G^X}{\bar{T}} = \frac{\bar{Y}}{\bar{T}}RS \qquad \text{and} \qquad \frac{W_Z^{Y|X} - (1-e)W_Z^X}{\bar{T}} = \frac{\bar{Y}}{\bar{T}}RSI.$$

5. An application to real tax data

5.1. The data and the micro-simulation model

To compare the Gini- and Zenga-based approaches applied to a real-world tax system, we use a static micro-simulation model of the Italian personal income tax, updated to the 2014 fiscal year. The model considers the most important taxes and contributions in the Italian fiscal system. It has been developed by Pellegrino (2007) about 10 years ago using the statistical software Stata, and it is constantly updated to incorporate changes in the tax code.

Here we focus on the module of the microsimulation model concerning the personal income tax that is an updated version of the model described in Pellegrino et al. (2011). Technical details regarding the structure and main results of this version of the micro-simulation model can be found in Pellegrino et al. (2019). The model employs, as input, income data provided by the Bank of Italy (2015) Survey on Household Income and Wealth (hereafter SHIW). This survey collected information on individual and household post-tax income and wealth in 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 raw data contained in the original survey have first to be appropriately reworked in order to determine the post-tax income income subject to the personal income tax. Then it is possible to apply the net-to-gross procedure, following the methodology proposed in Immervoll and O'Donoghue (2001), and to consequently determine the pre-tax distribution.

Considering individual taxpayers, results concerning the gross income distribution, the distribution of all tax variables, and the overall tax revenue are very close to the official statistics of the Department of Finance, Ministry of Economy and Finance (2016). Moreover, inequality indices, both for the taxpayers and equivalent households, are also very close to the ones evaluated by the official micro-simulation model of the Italian Department of Finance (Di Nicola et al., 2015). Therefore, the selected instrument is suitable for our empirical analysis.

Finally, to conduct our study, individual nominal incomes have to be transformed into equivalent household incomes using a proper equivalence scale. We choose to adopt the equivalence scale given by the square root of the number of the household components.

5.2. Basic results on the Italian personal income tax

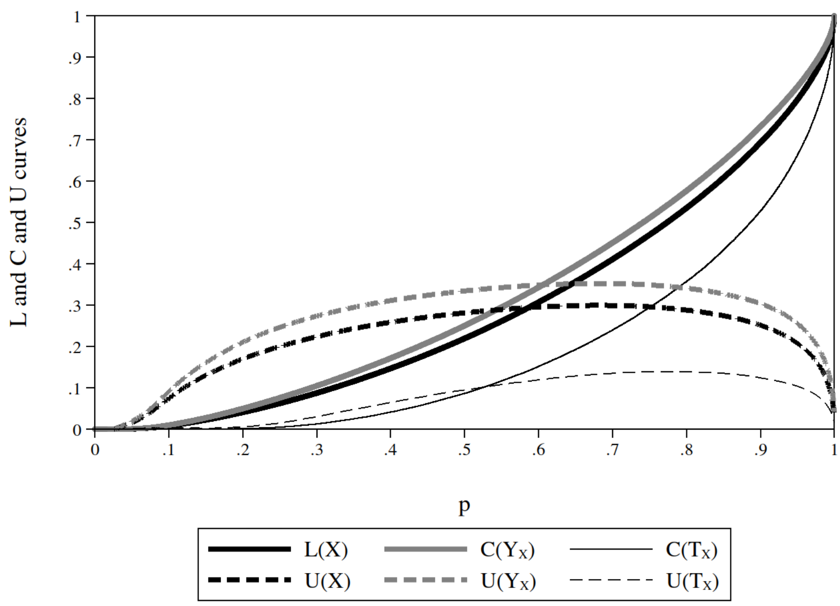
We begin our discussion by evaluating the inequality indices, presented in Section 4, using the data obtained from the micro-simulation model applied on the SHIW. We get $G(X) = 0.42058$, and by going back to the definition of the Gini coefficient in (4), this means that, before tax, on average, the mean income of the poorest groups is equal to 57.94% of the overall mean. By considering income data before tax, and interpreting the definition of the Zenga inequality index in (6), $I(X) = 0.77869$ indicate that by splitting the population in two complementary groups at each percentile and averaging over all ratios, the mean income of the poorest groups is 22.13% of the mean income of the richest groups.

Additionally, were the tax scheme proportional, the concentration coefficient for the tax liability $C(T_X)$ and the corresponding Gini coefficient $G(T)$ would be equal to the Gini coefficient for the pre-tax income $G(X)$, and, analogously, $I(T_X)$ and $I(T)$ would be equal to $I(X)$. Since the tax is progressive, we have instead $G(T) = 0.64179$, $C(T_X) = 0.63541$, thus $K = 0.21483$.

Similarly, $I(T) = 0.92849$, $I(T_X) = 0.92465$, so that $KI = 0.14596$. They measure the departure from proportionality of the tax. Moreover, were the tax scheme proportional, the concentration coefficient for post-tax incomes $C(Y_X)$ would be equal to the Gini coefficient for the pre-tax income. Since the tax is progressive, we have that $C(Y_X) < G(X)$, namely, $C(Y_X) = 0.37046$ (and $I(Y_X) = 0.73453 < I(X)$). Further, we get $RS = 0.05012$ to measure the reduction in the inequality due to progressive taxation. A similar interpretation is due to $RSI = 0.04416$. For what concerns the re-ranking (see Equations (13) and (14)), from previous coefficients follow $R(Y_X) = 0.00058$ and $RI = 0.00051$.

In the first part of the paper, we stressed that the value added of the Zenga approach can be mainly appreciated by focusing on the uniformity curve. Therefore, Figure 1 plots the Lorenz and Zenga curves for the pre-tax distribution and the concentration and the Zenga curves for the post-tax and the tax liability distributions.

Figure 1: Lorenz and Zenga U curves for X, Y, and T distributions

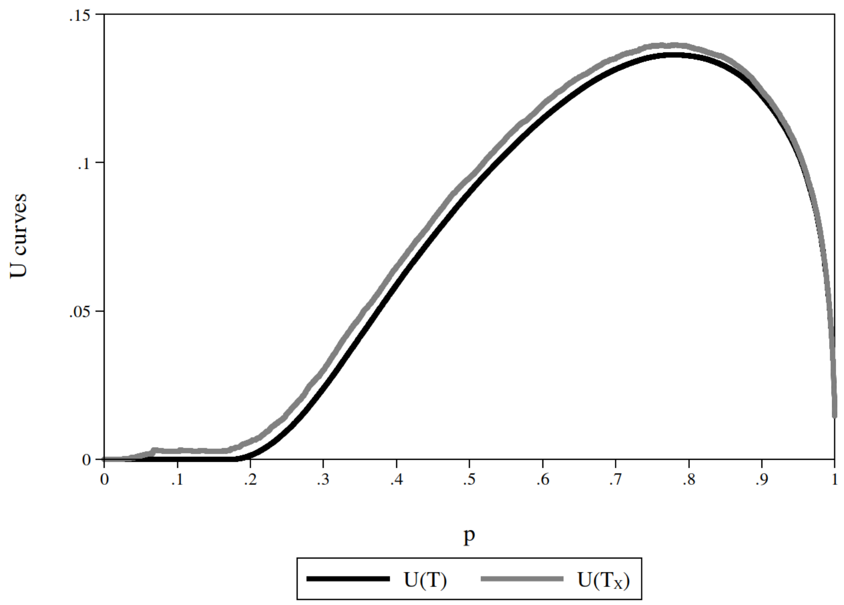


Looking at $U_X(p)$, we see that the bottom 25% of households earn a mean gross income that is equal only to 20.0% of the mean gross income of the top 75%; similarly, the bottom 50% earns a mean gross income equal to 28.2% of the one earned, on average, by the top half, with the corresponding percentages characterizing the bottom 75% and the bottom 99% being 29.6% and 11.6%, respectively.

This informative set complements the one derived by the Lorenz curve for pre-tax incomes, stating that the bottom 25% of households gets only 6% of the overall pre-tax incomes, the bottom 50% gets 22%, and so on. A similar interpretation is given by considering $U_{T_X}(p)$ and $U_{Y_X}(p)$. The tax liability distribution is more concentrated than the pre-tax income distribution; thus, $U_{T_X}(p)$ lies below the curve $U_X(p)$: the mean value of the tax liability paid by the bottom 25% of households is 1.5% of the mean value paid by the remaining 75% of households; similarly, the mean value of the tax liability paid by the bottom 50% of households is only 9.5% of the mean value paid by the remaining 50% of households, with the corresponding percentages of the bottom 75% and 99% being 13.9% and 5.7%, respectively. Conversely, the curve $U_{Y_X}(p)$ lies above the curve $U_X(p)$, since the relative position of households gets compressed in the transition from pre- to post-tax incomes.

The usefulness of the Zenga approach is apparent when studying the effect of progressive taxation in greater details. The graphical representation of the Lorenz and concentration curves for tax liability (here omitted) does not add information: the two curves are indistinguishable, as they lie approximately one above the other. In contrast, Figure 2 plots the corresponding curves $U_T(p)$ and $U_{T_X}(p)$ derived by the Zenga approach, and the difference between the two Zenga curves is apparent when the ordering differs.

Figure 2: The Zenga curves for T



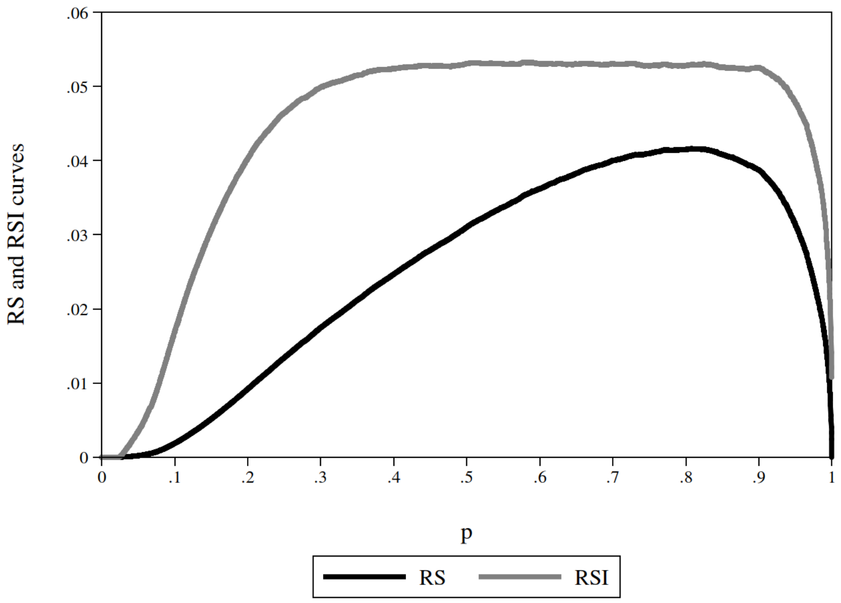
The re-ranking of taxes starts from very low values of the percentile (in the abscissa, about 0.05). It increases until 0.20, becomes constant up to 0.80, and then starts decreasing. Therefore, we can appreciate the effectiveness of adopting the Zenga approach to study the effects of the progressive taxation.

By construction, the Zenga-based approach allows the observation of such tiny differences, since $U_T(p)$ and $U_{T_X}(p)$ span from zero to 0.15, compared to the corresponding Gini-based approach, where the Lorenz and concentration curves always span from zero to 1.

Finally, Figures 3 and 4 plot the RS and K effects, comparing the two approaches. The standard Lorenz-based analysis, in both cases, generates curves (see their definitions in (19), (20), (21), and (22)) that monotonically increase up to the 85th percentile, and then decrease. In other words, the share of post-tax income accruing to each poorest portion of population, up to the 8th decile, is greater than the corresponding share of pre-tax incomes, because of progressive taxation.

The Zenga approach provides us with different information. The $RSI(p)$ effect, shown in Figure 3, increases up to the 40th percentile, then more or less becomes constant up to the 90th percentile. In particular, the curve $RSI(p)$ increases more sharply than the curve $RS(p)$ up to the 30th percentile, demonstrating that the poor benefit the most from progressive taxation; however, as the Italian tax system does not consider a negative income taxation, the absolute benefit depends on the available pre-tax income. Finally, starting from the 90th percentile, the $RSI(p)$ effect sharply decreases, since the top 10% of households pay a very high share of the overall tax revenue when compared to the tax revenue provided by the bottom 90% of the population, with this share increasing with income.

Figure 3: RS according to the Lorenz and Zenga approach



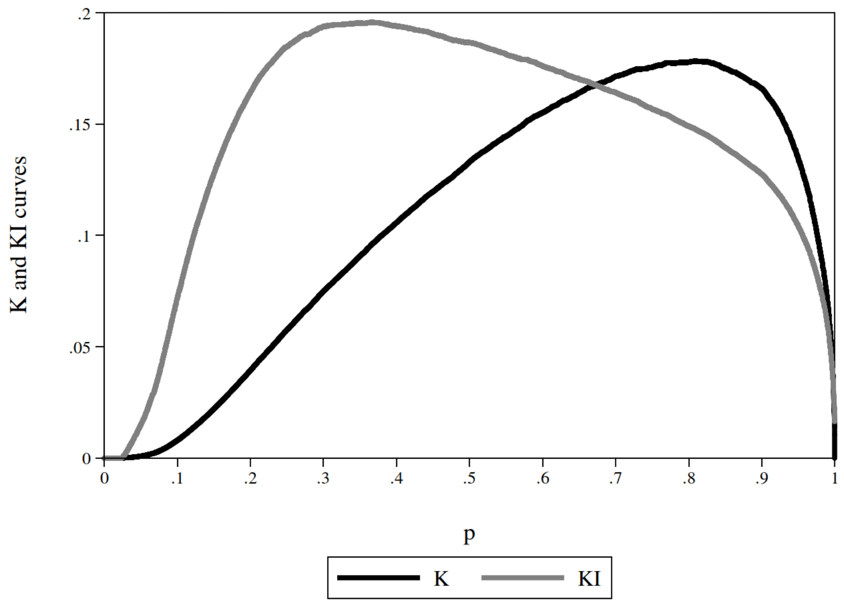
Moreover, $U_X(p)$ and $U_{Y_X}(p)$ increase up to the 68th percentile, with these increases being remarkably high up to the 40th percentile, underlining that, up to the 4th decile, the mean pre- and post-tax income of the poorest part of the distribution increase faster than the mean pre- and post-tax income of the richest part of the distribution. From the 40th percentile to the 70th percentile, this relationship remains, but loses its power; then from the 70th percentile, it is reversed, particularly from the 90th percentile, as observed before. On the contrary, the distance between $U_X(p)$ and $U_{T_X}(p)$ decreases from the 30th percentile onwards, before what is observed for the curve $RSI(p)$ (Figure 4). Evidently, by comparing the share of gross income and taxes accruing to each decile, it can be observed that their distance is increasing up to the 30th percentile, then it decreases, being even negative in the top part of the income distribution.

Given these observations, $U_X(p)$, $RSI(p)$, and $KI(p)$ place an emphasis on the part of the income distribution that benefits the most and the part that is harmed the most by progressive taxation (with respect to a proportional tax yielding the same revenue).

In addition, the shape of the curve $RSI(p)$ seems to show a pattern related to the shape of the average tax rate in each percentile, with the latter being zero up to the 18th percentile, then it sharply increases up to the 30th percentile. Afterwards, it continues increasing with less intensity up to the 90th percentile. Finally, in the top part of the income distribution, the average tax rate is very high.

To understand this relation, note that, according to 19, the curves $RS(p)$ and $K(p)$ are related by the fixed coefficient $\frac{\alpha}{1-\alpha}$. However, according to 23, the curves $RSI(p)$ and $KI(p)$ are related by the factor $\frac{\alpha^+(p)}{1-\alpha^+(p)}$, which depends on the percentile, and allows to determine RSI from KI.

Figure 4: K according to the Lorenz and Zenga approach



Focusing on this remark, Figure 4 compares the two Kakwani effects. As mentioned before, by adopting the Lorenz approach, the progressivity effect increases up to the 90th percentile, while according to the Zenga curve, it increases only up to the 30th percentile.

The shape of the curve $KI(p)$ is very intuitive, as in the bottom part of the income distribution, $U_X(p)$ is strictly increasing, while about 20% of the households pay no taxes, so that the curve $U_{T_X}(p)$ is zero. Therefore, their difference, that is $KI(p)$, strictly increases once all households with nil tax liability have been considered. Starting from the 40th percentile onwards, $U_{T_X}(p)$ increases at a higher rate than $U_X(p)$, so that the curve $KI(p)$ decreases.

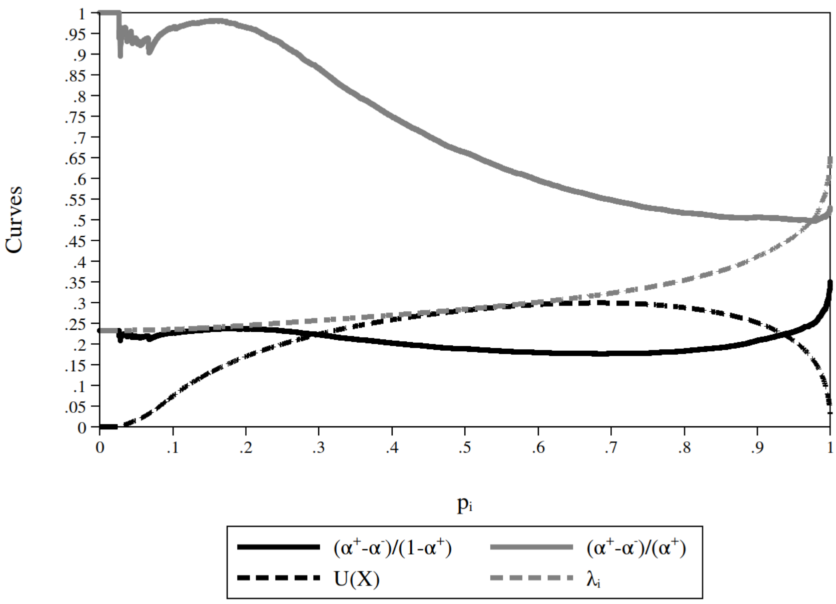
Following the discussion presented in Subsection 4.2, we now observe the behaviour of the most important components of $RS(p)$ and $RSI(p)$, in particular, in the range $0.4 < p < 0.9$.

Recalling (21) and (22), Figure 5 shows all the components that participate in the definition of the RSI and KI curves. The uniformity curve $U_X(p)$ (dashed black line) and $\frac{\alpha^+(p)-\alpha^-(p)}{1-\alpha^+(p)}$ (solid black line) contribute to define $RSI(p)$, while the uniformity curve $U_X(p)$ and $\frac{\alpha^+(p)-\alpha^-(p)}{\alpha^+(p)}$ (solid gray line) contribute to define $KI(p)$. Recalling (15) and (23), we get $\lambda_i = \frac{\alpha^+(p)}{1-\alpha^+(p)}$ (dashed gray line).

For $0.3 < p < 0.9$, the dashed and solid black lines of Figure 5 are almost symmetric, and the product of their ordinates is almost constant, explaining the constant gray line ($RSI(p)$) of Figure 3. Conversely, for $0 < p < 0.3$, the product of the ordinates of the solid and dashed black lines is increasing, whereas the opposite happens for $p > 0.9$.

The shapes (here omitted) of the curves $\alpha^+(p) - \alpha^-(p)$ and $\frac{\alpha^+(p)-\alpha^-(p)}{1-\alpha^+(p)}$ are very close to each other, with $\alpha^-(p)$ growing approximately at a constant rate in the range $0.4 < p < 0.9$. In other words, the curve $RSI(p)$ shows that the progressivity of the tax is not expansionist for this portion of population. Detailed information on the tax progressivity can be observed analysing the $KI(p)$ curve (gray line of Figure 4). Its ordinates are given by the product of the ordinates of the dashed black line and the solid gray line of Figure 5), that first increases (up to the 30th percentile), and then decreases.

Figure 5: The contributions to RSI



5.3. Social welfare implications of taxation

Having described results according to the usual statistical framework, we discuss here the social welfare implications of both Gini and Zenga approach (see Subparagraph 4.3).

Focusing on Eq. (5), the overall social welfare when pre-tax incomes are considered is $W_G^X = 12,477.86$. The overall social welfare for post-tax incomes is lower, since here we focus our attention only on the (personal income) taxation side, without considering the welfare gains due to public expenditures. In particular, the overall social welfare when post-tax incomes are considered, according to the post-tax order, is $W_G^Y = 10,982.50$; the overall social welfare when post-tax incomes are considered, according to the pre-tax order, is $W_G^{Y|X} = 10,992.63$. In the transition from the pre-to the post-tax incomes, a social loss of about 11.9% is registered: $W_G^Y - W_G^X = -1,495.36$.

Decomposing this overall value can give more details on the welfare loss due to taxation (Eq. (24)):

$$W_G^Y - W_G^X = -eW_G^X + \left(W_G^{Y|X} - (1-e)W_G^X\right) + H_G^A.$$

We get[9]

$$-1,495.36 = -(2,360.38) + (875.16) + (-10.13).$$

This means that, according to the Gini social welfare function, were all income units are asked to pay the same amount of taxes yielding the same tax revenue observed in the actual tax system, the per capita social welfare loss would be $-2,360 = -eW_G^X$ euro. However, the tax system is progressive, so that the tax progressivity of the Italian personal income tax, measured by $W_G^{Y|X} -$

[9]By letting $-1,495.36$ be equal to 100, the first term is about 0.68, the second one is -157.85 and the third one is -58.53.

$(1-e)W_G^X$, determines a per capita increase of social welfare by about 875 euro; this increase only partially overcomes the negative effect due to $-eW_G^X$. The application of the tax also determines inefficiencies; in particular, in the transition from the pre- to the post-tax, income reranking occurs: this unpleasant outcome is measured by the welfare loss due to horizontal inequity H_G^A, that is very small, about 10 euro per capita.

A similar picture emerges whenever the Zenga approach is considered. Focusing on Eq. (10), the overall social welfare when pre-tax incomes are considered is $W_Z^X = 4,765.66$, clearly lower than W_G^X. The welfare level according to the Zenga approach is lower than the one evaluated according to the Gini methodology in our empirical exercise: in the latter the weights $2(1-p)$ are linearly decreasing from 2 to zero when p increases, whilst in the former the values of the weights $w_Z(p) = \frac{-ln(p+p-1)}{(1-L_F(p))^2}$ are higher than $w_G(p) = 2(1-p)$ for $p < 0.059$, and they are always lower elsewhere.

Similarly, the overall social welfare for post-tax incomes, according to the post-tax order, is lower than both W_Z^X and W_G^X, and it is equal to $4,626.34$. Finally, the overall social welfare when post-tax incomes are considered, according to the pre-tax order, is instead $W_Z^{Y|X} = 4,635.18$.

In the transition from the pre-to the post-tax incomes, an overall social loss (see Eq. (25)) of about 2.9% is registered: $W_Z^Y - W_Z^X = -139.32 = H_Z^A - eW_Z^X + \left(W_Z^{Y|X} - (1-e)W_Z^X\right)$. According to this decomposition, $-eW_Z^X = -901.50$, $W_Z^{Y|X} - (1-e)W_Z^X = 771.02$, and $H_Z^A = -8.85$. According to the Zenga social welfare function, the tax progressivity of the tax determines a per capita increase of social welfare by about 771 euro; like in the Gini approach, this increase only partially overcomes the other negative effects. The loss of social welfare that is due to the application of a proportional tax yielding the same tax revenue is equal to 901 euro per capita. The welfare loss due to the horizontal inequity is small also in this case, about 9 euro per capita.

We do not discuss results according to the relative decomposition: as mentioned above, the measures of that decomposition are simply obtained by dividing each term of the absolute decomposition by the average tax amount $\bar{T} = 4,073.68$.

6. Concluding remarks

In this paper, we have applied a recently proposed approach to measure inequality, namely, the Zenga index and curve, to study the degree of progressivity and the redistributive and re-ranking effects of taxation. The novel approach to inequality proposed by Zenga (2007) is based on contrasting the economic conditions of opposite and exhaustive parts of the population.

After analyzing how the Zenga uniformity curve reacts with respect to variations in the Lorenz curve, we have derived the first original result of our research: the elasticity of the Zenga uniformity curve exceeds the elasticity of the Lorenz curve, and increases with the population percentile.

Therefore, following the existing tax literature, we have replicated the most important curves and the corresponding tax indices, namely, the Reynolds-Smolensky, the Kakwani, and the Atkinson-Plotnick-Kakwani indices, adopting the new paradigm. Our proposal is motivated by the fact that valid information for policy makers is obtained by comparing the degree of progressivity and the redistributive and the re-ranking effects of a personal income tax system observed in the poorest part of the population, with the corresponding degree and effects occurring in the richest part. Furthermore, along the lines of the well-known properties of the three aforementioned indices, we have derived the corresponding properties that hold true for the newly introduced measures, and we have shown how social welfare gains and losses due to taxation can be assessed in the new approach.

We then discussed the strengths and weaknesses of our approach, by comparing the information conveyed by the classical indices and the new ones through applying it to the Italian tax system, based on a micro-simulation model of the personal income tax, updated to the 2014 fiscal year. We have employed, as input data for the model, the Bank of Italy (2015) Survey on Household Income and Wealth. We have shown that the new curves provide an insight into information that could be hidden (or at least diminished) in the cumulative approach intrinsic to the Lorenz curve. In light of the obtained results, the analysis of the effects of a tax system has been enriched by the considerations derived from the new approach.

Acknowledgments

We would like to thank Ivica Urban and three anonymous referees for their useful comments that helped us improve the paper. Usual disclaimers apply.

References

AABERGE, R. (2000): "Characterizations of Lorenz curves and income distributions," *Social Choice and Welfare*, 17, 639–653.

ANTAL, E., M. LANGEL, AND Y. TILLÉ (2011): "Variance estimation of inequality indices in complex sampling designs," *Proceedings of the* 58^{th} *World Statistical Congress.*

ARCAGNI, A. AND F. PORRO (2014): "The Graphical Representation of Inequality," *Revista Colombiana de Estadistica*, 37, 419–436.

ARCAGNI, A. AND M. ZENGA (2013): "Application of Zenga's distribution to a panel survey on household incomes of European Member States," *Statistica & Applicazioni*, 11, 79–102.

ATKINSON, A. B. (1970): "On the Measure of Inequality," *Journal of Economic Theory*, 2, 244–263.

——— (1980): "Horizontal equity and the distribution of the tax burden," in *The economics of taxation*, ed. by H. J. Aaron and M. J. Boskins, Washington: The Brookings Institution, chap. 1, 3–18.

ATKINSON, A. B. AND F. BOURGUIGNON (1987): "Income Distribution and Differences in Needs," in *Arrow and the Foundations of the Theory of Economic Policy*, ed. by G. R. Feiwel, London: Palgrave Macmillan, chap. 12, 350–369.

BANK OF ITALY (2015): "Household income and wealth in 2014, Supplements to the Statistical Bulletin, Year XXV (New Series), No. 64," .

DAVYDOV, Y. AND F. GRESELIN (2018): "Comparisons between poorest and richest to measure inequality," *Sociological Methods & Research.*

Department of Finance, Ministry of Economy and Finance (2016): "Statistical Reports," .

Di Nicola, F., G. Mongelli, and S. Pellegrino (2015): "The static microsimulation model of the Italian Department of Finance: Structure and first results regarding income and housing taxation," *Economia Pubblica*, 2, 125–157.

Donaldson, D. and J. A. Weymark (1980): "A single-parameter generalization of the Gini indices of inequality," *Journal of Economic Theory*, 22, 67–86.

Gastwirth, J. L. (2014): "Median-based measures of inequality: Reassessing the increase in income inequality in the US and Sweden," *Statistical Journal of the IAOS*, 30, 311–320.

——— (2016): "Measures of Economic Inequality Focusing on the Status of the Lower and Middle Income Groups," *Statistics and Public Policy*, 3, 1–9.

Gini, C. (1912): *Variabilità e Mutuabilità. Contributo allo Studio delle Distribuzioni e delle Relazioni Statistiche*, C. Cuppini, Bologna.

——— (1914): "Sulla misura della concentrazione e della variabilità dei caratteri," *Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti*, 73, 1203–1248.

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

Greselin, F., L. Pasquazzi, and R. Zitikis (2010): "Zenga's New Index of Economic Inequality, Its Estimation, and an Analysis of Incomes in Italy," *Journal of Probability and Statistics*, Special Issue on "Actuarial and Financial Risks: Models, Statistical Inference, and Case Studies".

——— (2013): "Contrasting the Gini and Zenga Indices of Economic Inequality," *Journal of Applied Statistics*, 40, 282–297.

——— (2014): "Heavy Tailed Capital Incomes: Zenga Index, Statistical Inference, and ECHP Data Analysis," *Extremes*, 17, 127–155.

Greselin, F., S. Pellegrino, and A. Vernizzi (2020): "The social welfare implications of the Zenga index," *Arkiv.*

Greselin, F., M. L. Puri, and R. Zitikis (2009): "L-functions, Processes, and Statistics in Measuring Economic Inequality and Actuarial Risks," *Statistics and Its Interface*, 2, 227–245.

Guillaud, E., M. Olckers, and M. Zemmour (2019): "Four Levers of Redistribution: The Impact of Tax and Transfer Systems on Inequality Reduction," *Review of Income and Wealth*, 66, 444–466.

Immervoll, H. and C. O'Donoghue (2001): "Imputation of gross amounts from net incomes in household surveys. An applicationn using EUROMOD," *EUROMOD Working Paper, No. No. EM1-01.*

Jakobsson, U. (1976): "On the Measurement of the Degree of Progression," *Journal of Public Economics*, 5, 161–168.

Jasso, G. (1982): "Measuring inequality: using the geometric mean/arithmetic mean ratio," *Sociological Methods & Research*, 10, 303–326.

——— (2018): "Anything Lorenz Curves Can Do, Top Shares Can Do: Assessing the TopBot Family of Inequality Measures," *Sociological Methods & Research.*

Jedrzejczak, A. and K. Trzcinska (2019): "Application of the Zenga distribution to the analysis of household income in Poland by socio-economic group," *Statistica & Applicazioni*, 123–140.

Kakwani, N. (1977a): "Applications of Lorenz curves in economic analysis," *Econometrica*, 719–727.

——— (1977b): "Measurement of tax progressivity: an international comparison," *Economic Journal*, 87, 71–80.

——— (1984): "On the measurement of tax progressivity and redistributive effects of taxes with applications to horizontal and vertical equity," in *Economics inequality, measurement and policy*, ed. by G. F. Rhodes and R. L. Basmann, 149–168.

Kakwani, N. and H. Son (2019): "Normative Measures of Tax Progressivity An International Comparison," .

Lambert, P. J. (2001): *The distribution and redistribution of income*, Manchester University Press.

Langel, M. and Y. Tillé (2012): "Inference by Linearization for Zenga's New Inequality Index: a Comparison with the Gini Index," *Metrika*, 75, 1093–1110.

Lorenz, O. (1905): "Methods of Measuring the Concentration of Wealth," *Publications of the American Statistical Association*, 9, 209–219.

Maffenini, W. and M. Polisicchio (2014): "The potential of the I(p) inequality curve in the analysis of empirical distributions," *Statistica & Applicazioni*, 12, 63–85.

Musgrave, R. A. and T. Thin (1948): "Income Tax Progression, 1929-48," *Journal of Political Economy*, 56, 498–514.

Mussini, M. and M. Zenga (2013): "A longitudinal decomposition of Zenga's new inequality index," *Statistica & Applicazioni*, 11, 63–77.

Osberg, L. (2017): "On the Limitations of Some Current Usages of the Gini Index," *Review of Income and Wealth*, 63, 574–584.

Pasquazzi, L. and M. Zenga (2018): "Components of Gini, Bonferroni, and Zenga Inequality Indexes for EU Income Data," *Journal of Official Statistics*, 34, 149–180.

Pellegrino, S. (2007): "Il Modello di microsimulazione IRPEF 2004," Società Italiana di Economia Pubblica – SIEP, 2007, WP No. 583.

Pellegrino, S., P. Massimiliano, and T. Gilberto (2011): "Developing a static microsimulation model for the analysis of housing taxation in Italy," *The International Journal of Microsimulation*, 4, 73–85.

Pellegrino, S., G. Perboli, and G. Squillero (2019): "Balancing the equity-efficiency trade-off in personal income taxation: an evolutionary approach," *Economia Politica*, 36, 37–64.

Pietra, G. (1915): "Delle relazioni tra gli indici di variabilità," *Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti*, 74, 775–804.

Piketty, T. (2013): *Le capital au XXIe siècle*, Le Seuil.

Plotnick, R. (1981): "A Measure of Horizontal Inequity," *The Review of Economics and Statistics*, 63, 283–288.

Polisicchio, M. (2008): "The continuous random variable with uniform point inequality measure I(p)," *Statistica & Applicazioni*, 6, 137–151.

Polisicchio, M. and F. Porro (2009): "A comparison between Lorenz L(p) curve and Zenga I(p) curve," *Statistica Applicata*, 21, 289–301.

Radaelli, P. (2008): "A Subgroups Decomposition of Zenga's Uniformity and Inequality Indexes," *Statistica & Applicazioni*, 6, 117–136.

——— (2010): "On the Decomposition by Subgroups of the Gini Index and Zenga's Uniformity and Inequality Indexes," *International Statistical Review*, 78, 81–101.

Reynolds, M. and E. Smolensky (1977): *Public expenditures, taxes and the distribution of income: the United States, 1950, 1961, 1970*, New York Academic Press.

Sen, A. (1974): "Informational bases of alternative welfare approaches: aggregation and income distribution," *Journal of Public Economics*, 3, 387–403.

Shorrocks, A. F. (1983): "Ranking Income Distributions," *Economica*, 50, 3–17.

Son, H. H. (2013): *Equity and well-being: measurement and policy practice*, Routledge.

Vergnaud, J.-C. (1997): "Analysis of risk in a non-expected utility framework and application to the optimality of the deductible," *Revue Finance*, 18, p155–167.

Yitzhaki, S. (1983): "On an Extension of the Gini Inequality Index," *International Economic Review*, 24, 617–628.

Zenga, M. (2007): "Inequality Curve and Inequality Index Based on the Ratios Between Lower and Upper Arithmetic Means," *Statistica & Applicazioni*, 5, 3–27.

Zenga, M., P. Radaelli, and M. Zenga (2012): "Decomposition of the Zenga's Inequality Index by Sources," *Statistica & Applicazioni*, 9, 3–34.

Appendix A: Technical proofs

Proof of the properties of the Zenga index

Property i) follows from the fact that if $c > 0$, then $F^{-1}_{c\,Z}(p) = c\,F^{-1}_Z(p)$ and $\bar{\mu}_{c\,Z}(p)/\mu^+_Z(p) = \mu^-_Z(p)/\mu^+_Z(p)$, for every $c \in (0,1)$.

Property ii) is obtained from the linearity of the mean, and the fact that for every $p > 0$, we have the bound $\mu_Z^+(p) \geq \mu_Z^-(p)$; thus, $\frac{\mu_{Z+c}^-(p)}{\mu_{Z+c}^+(p)} = \frac{\mu_Z^-(p)+c}{\mu_Z^+(p)+c} \geq \frac{\bar{\mu}_Z(p)}{\mu_Z^+(p)}$ for all $c > 0$. The latter bound is equivalent to $I_{Z+c}(p) \leq I_Z(p)$ for every $p \in (0,1)$, establishing the bound $I_{Z+c} \leq I_Z$.

Property iii) can be easily shown using the relationship between $I_F(p)$ and $L_F(p)$ given in Zenga (2007), that is, $I_F(p) = \frac{p-L_F(p)}{p(1-L_F(p))}$.

Finally, property iv) follows from recalling that Z is less unequal than Y in the Pigou-Dalton sense (see Vergnaud (1997)), denoted by $Z \leq_{PD} Y$, if and only if $\mu_Z = \mu_Y$ and $Z \leq_L Y$.

Thus, the Pigou-Dalton property is now obvious. □

Proof of the relationship between the Kakwani progressivity index and the Reynolds-Smolensky index, reframed within the Zenga approach

To verify Eq. (15) we start from the equality

$$\frac{\mu_X^+(p) - \mu_X^-(p)}{\mu_X^+(p)} - \frac{\mu_{Y_X}^+(p) - \mu_{Y_X}^-(p)}{\mu_{Y_X}^+(p)} = \frac{\mu_{T_X}^+(p) - \mu_{T_X}^-(p)}{\mu_{T_X}^+(p)}, \tag{26}$$

which is immediately verified. From Eq. (26) we get:

$$\left(\frac{\mu_X^+(p) - \mu_X^-(p)}{\mu_X^+(p)} - \frac{\mu_{Y_X}^+(p) - \mu_{Y_X}^-(p)}{\mu_{Y_X}^+(p)}\right)\frac{\mu_X^+(p)}{\mu_{Y_X}^+(p)} = \frac{\mu_{T_X}^+(p) - \mu_{T_X}^-(p)}{\mu_{T_X}^+(p)}\ \frac{\mu_X^+(p)}{\mu_{Y_X}^+(p)},$$

which, after trivial simplifications, becomes

$$\frac{\mu_X^+(p) - \mu_X^-(p)}{\mu_{Y_X}^+(p)} - \frac{\mu_{Y_X}^+(p) - \mu_{Y_X}^-(p)}{\mu_{Y_X}^+(p)} = \frac{\mu_{T_X}^+(p) - \mu_{T_X}^-(p)}{\mu_{T_X}^+(p)}\ \frac{\mu_{T_X}^+(p)}{\mu_{Y_X}^+(p)}. \tag{27}$$

By adding to the l.h.s. of Eq. (27) the quantity

$$\frac{\mu_X^+(p) - \mu_X^-(p)}{\mu_X^+(p)} - \frac{\mu_X^+(p) - \mu_X^-(p)}{\mu_{Y_X}^+(p)}, \tag{28}$$

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

$$-\frac{\mu_X^+(p) - \mu_X^-(p)}{\mu_X^+(p)}\ \frac{\mu_X^+(p) - \mu_{Y_X}^-(p)}{\mu_{Y_X}^+(p)} = -\frac{\mu_X^+(p) - \mu_X^-(p)}{\mu_X^+(p)}\ \frac{\mu_{T_X}^+(p)}{\mu_{Y_X}^+(p)},$$

then

$$\frac{\mu_X^+(p) - \mu_X^-(p)}{\mu_X^+(p)} - \frac{\mu_{Y_X}^+(p) - \mu_{Y_X}^-}{\mu_{Y_X}^+(p)} =$$
$$\left(\frac{\mu_{T_X}^+(p) - \mu_{T_X}^-(p)}{\mu_{T_X}^+(p)} - \frac{\mu_X^+(p) - \mu_X^-(p)}{\mu_X^+(p)}\right)\frac{\mu_{T_X}^+(p)}{\mu_{Y_X}^+(p)},$$

which is Eq. (15) with $\lambda(p) = \frac{\mu^+_{T_X}(p)}{\mu^+_{Y_X}(p)}$. $\square$

Proof of the relationship between the Lorenz and the uniformity curves for the lowest and highest percentiles

We firstly show that $U(p) \geq L(p)$ for all $p \leq \mu_Z/\mu^+_Z(p)$. We have

$$\begin{aligned} L_Z(p) &= \frac{p\ \mu^-_Z(p)}{\mu_Z} \\ &\leq \frac{\mu^-_Z(p)}{\frac{\mu_Z}{p}} = \frac{\mu^-_Z(p)}{\mu^+_Z(p)} = U_Z(p). \end{aligned}$$

Note that the function $h(p) = \mu_Z/\mu^+_Z(p)$ takes values in $[1, \mu_Z/max(Z)]$ and is a decreasing function. Now, for $p \geq \mu_Z/\mu^+_Z(p)$, we have the converse $U_Z(p) \leq L_Z(p)$.

$\square$

Appendix B: Empirical estimates of the Gini and Zenga indices

In sections 2 and 3 of the paper, we have presented, in a theoretical setting, the Lorenz and Zenga curves and the Gini and Zenga indices as functionals defined on the space of the distribution functions. Naturally, we do not know the analytical form of the income distribution for the population; thus, we must rely on survey data for inferring our knowledge of inequality. Therefore, we need formulas for estimating the aforementioned curves and indices on samples of real data, which is the aim of the present Section, where we provide the empirical counterparts of the previous quantities. Let us now suppose that we have data coming from a survey of n households from a population, where units included in the sample have unequal probabilities of selection. To correct for this possible source of bias, each observation is weighted using the inverse of the sampling fraction (sampling weight) adjusted for the non-response mechanism (non-response weight) and often incorporates auxiliary information about the population (post-stratification). Let $z_1, \ldots, z_n$ be the ranked observations of a quantitative variable Z (e.g. the household gross income, tax, or net income), and let w_i denote the corresponding survey weights w_i. Therefore, the Zenga index compares the average of Z, from the first household to the i^{th} poorer one, as

$$M^-_i(Z) = \frac{\sum_{j=1}^{i} z_j w_j}{\sum_{j=1}^{i} w_j}$$

with the average of the remaining $n - i$ households, being

$$M^+_i(Z) = \frac{\sum_{j=i+1}^{n} z_j w_j}{\sum_{j=i+1}^{n} w_j},$$

then the ratio

$$I_i(Z) = \frac{M_i^+(Z) - M_i^-(Z)}{M_i^+(Z)}$$

evaluates the inequality at each empirical percentile $p_i = \sum_{j=1}^{i} w_j / \sum_{j=1}^{n} w_j$ of the distribution. Points $\{(p_i, I_i(Z)); i = 1, \ldots, n-1\}$ are connected by adjacent segments, to obtain a curve in $[0,1] \times [0,1]$ as for the empirical Lorenz curve. If all income earners receive the same income, all ratios are equal, while their variability is a direct function of inequality. We will consider also the complementary curve $(p_i, U_i(Z))$, called the uniformity curve, defined by

$$U_i(Z) = 1 - I_i(Z) = \frac{M_i^-(Z)}{M_i^+(Z),}$$

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, for $i = 1$ to $n-1$. Moreover, they are not constrained to assuming a fixed value at their end points, different from $L_i(Z) = \frac{\sum_{j=1}^{i} z_j w_j}{\sum_{j=1}^{n} z_j w_j}$, which always begins in $(0,0)$ and ends in $(1,1)$.

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

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

and

$$I_i(Z) = 1 - \frac{1-p_i}{p_i} \frac{L_i(Z)}{1 - L_i(Z)}.$$

Finally, we recall the empirical estimators $\widehat{I}(Z)$ for the Zenga index (Greselin et al., 2010):

$$\widehat{I}(Z) = \frac{1}{n-1} \sum_{i=1}^{n-1} \frac{M_i^+(Z) - M_i^-(Z)}{M_i^+(Z)} = 1 - \widehat{U}(Z)$$

and $\widehat{G}(Z)$ for the Gini index, respectively

$$\widehat{G}(Z) = \frac{1}{n-1} \sum_{i=1}^{n-1} \frac{\overline{Z} - M_i^-(Z)}{\overline{Z}}, \quad \text{where} \quad \overline{Z} = \frac{\sum_{j=1}^{n} z_j w_j}{\sum_{j=1}^{n} w_j}.$$

We remark that the weighted formulas have been employed in all our calculations on sample data in Section 5.