Mauro Mussini
Biancamaria Zavanella[1]

Choosing the Bandwidth for Decomposing the Redistributive Effect: Evidence from Milan Using AMeRIcA Data

1. Introduction

The literature offers a number of approaches to decompose the total redistributive effect of a tax system into three different components: a vertical effect, an horizontal effect and a reranking effect. In this frame, given the sparseness of exact equals in real world data sets, it arises the issue to suggest a procedure to define the close equals groups optimally in terms of class width. Vernizzi and Pellegrino (2007) suggest a criterion to choose a convenient bandwidth in defining the close equals groups and the aim of this paper is to verify either the coherence or the validity of this criterion. In order to pursue this goal, we analyse the results of the redistributive effect decomposition by comparing the models proposed by Aronson, Johnson and Lambert (1994), van de Ven, Creedy and Lambert (2001) and Urban and Lambert (2008).

As it is known, Aronson, Johnson and Lambert (1994), henceforth AJL, suggest a decomposition of the total redistributive effect based on the assumption that taxpayers are split into groups formed by exact pre-tax equals. In 2001, van de Ven, Creedy and Lambert, hereafter VCL, note that exact pre-tax equals are rarely observed in survey data, so they consider groups of close equals for de-

---

composing the redistributive effect. In Urban and Lambert (2008), henceforth UL, it is shown that VCL decomposition method does not capture entire reranking effect when reranking occurs within the groups of close equals or among group mean incomes. UL recall the approaches to decompose the redistributive effect presented in the previous works and introduce both a modifications to the original AJL decomposition model when dealing with groups of pre-tax close equals income receivers and a new decomposition of the redistributive effect. This last decomposition captures all forms of reranking.

As said above, Vernizzi and Pellegrino (2007), henceforth VP, suggest a criterion to choose a convenient bandwidth in defining the close equals groups. This criterion is shareable considering together all the quoted decompositions of the redistributive effect and the convenient bandwidth can be chosen analysing the behaviour of the vertical effects obtained by applying the alternative decomposition models.

Using 2004 Italian SHIW[2] data set, VP confine their analysis to bandwidths which ranges from 10 euro to 3000 euro: within this bandwidth range the three alternative measures of vertical effect are higher than the actual redistributive effect. In our analysis we extend the bandwidth range over the limit considered by VP, in order to verify if VP criterion leads to a proper choice even when the investigation is not limited to bandwidths which ensure vertical effects greater than the actual redistributive effect. Moreover in order to investigate if VP results may depend on their particular data set; we base our empirical analysis on AMeRIcA data set, which includes all taxpayers resident in the Milanese area for 2003.

Our empirical analysis shows that for the chosen bandwidth the three alternative measures of vertical effects explain a potential redistribution. This is the main result of the paper. Moreover the analysis of the behaviour of the different redistributive effect components allows us to make more clear to the effect of the income distribution skewness on determining the horizontal effect.

The paper is organized as follows. Section 2 outlines the decomposition methods introduced in the previous literature. Section 3 discusses literature suggestions about the choice of the optimal bandwidth, matching them with our empirical evidence which is extensively reported and discussed in section 4. Section 5 is devoted to final remarks.

---

[2] SHIW is the Bank of Italy survey on households incomes and wealth.

2. Preliminaries

Let G_y and G_x be the Gini coefficients associated with pre-tax and post-tax distributions respectively. The difference $RE = G_y - G_x$ measures the redistributive effect of taxation. The pre-tax Gini coefficient, G_y, can be decomposed into three terms (Dagum, 1997)

$$G_y = G_y^b + G_y^w + G_y^t, \tag{1}$$

where G_y^b is the pre-tax between groups Gini coefficient, G_y^w is the pre-tax within groups Gini coefficient and G_y^t is the component due to the presence of overlapping among groups.

When the population of taxpayers is divided into groups which contains exact pre-tax equals occurs that $G_y^w = 0$ and $G_y^t = 0$. Then the decomposition of the pre-tax Gini coefficient across groups becomes

$$G_y = G_y^b. \tag{2}$$

Moving from pre-tax to post-tax income distribution, net incomes may diverge within the same group of pre-tax equals; consequently within groups inequality component becomes different from zero. Moreover post-tax incomes of individuals belonging to different groups may overlap. If this is the case, either the within group term or the overlapping terms becomes different from zero, so that G_x should be decomposed as in (1)

$$G_x = G_x^b + G_x^w + G_x^t. \tag{3}$$

Taking now into account the decomposition for G_y and G_x as given in (2) and (3), according to AJL (1994) suggestion, we obtain

$$\begin{aligned} RE &= G_y - G_x^b - G_x^w - G_x^t, \\ RE &= V - H - R^{AJL}. \end{aligned} \tag{4}$$

AJL call $V = G_y - G_x^b$ vertical effect: it measures the redistribution that would occur if pre-tax equals are treated equally. In (4) the term $H = G_x^w$ represents the fall of the potential redistribution due to the different taxation of equals, it involves taxpayers that perceive the same level of pre-tax income, so it is termed horizontal effect. The term $G_x^t = R^{AJL}$ provides the loss of the redis-

tributive effect arising from the difference in pre-tax and post-tax ranking of taxpayers. In (4) framework, R^{AJL} coincides with the Atkinson-Plotnick-Kakwani index of reranking, R^{APK}, which measures the reranking effect moving from pre-tax to post-tax income distribution when income units are ungrouped (Atkinson 1980, Plotnick 1981, Kakwani, 1984).

Due to the sparseness of exact pre-tax equals in survey data sets, VCL adapt the decomposition model described above when groups are formed by pre-tax close equals rather then by exact equals. When considering groups of close equals, one assume the presence of inequality within groups before taxation then the pre-tax Gini coefficient does not reduce to the between component but it is given by the sum of the Gini between and within components (G_y^b and G_y^w).

VCL suggest the following redistributive effect decomposition

$$RE = G_y^b + G_y^w - G_x^b - G_x^w - G_x^t,$$
$$RE = V^{VCL} - H^{VCL} - R^{AJL}, \tag{5}$$

where $V^{VCL} = \left(G_y^b - G_x^b\right), H^{VCL} = \left(G_x^w - G_y^w\right)$ and $R^{AJL} = G_x^t$.

Both AJL and VCL approaches assume that reranking does not occur among group mean incomes and within group orderings. If this is not the case, R^{AJL} does not measure the entire reranking which results from taxation, but only reranking which involves income units belonging to distinct groups. However UL, considering the empirical evidence, notice that these forms of reranking are not rare in micro data sets. To solve this problem they introduce a decomposition method that capture all forms of reranking.

Firstly UL adapt the original AJL decomposition model when dealing with close equals groups. They consider an artificial tax system which treats proportionally taxpayers within each group. According to this tax system, the group average tax rate is applied to pre-tax incomes within each group of close equals, in order to maintain the same within group inequality after and before tax[3]. However the smoothed within group Gini coefficient $G_x^{sw} = \sum_k a_{k,x} G_{k,y}$ is generally different from the pre-tax Gini coefficient $G_y^w = \sum_k a_{k,y} G_{k,y}$, because the generic post-tax weight $a_{k,x}$ may differ from the corresponding pre-tax one, $a_{k,y}$

---

[3] This method produces a smoothing effect within groups of close equals that in the case of exact equals groups is obtained when each net income is replaced by the respective group mean income after tax. Hence, maintaining a proportional tax rate within groups, the post-tax Gini coefficient referred to each group remains equal to the pre-tax one.

(see VP, 2007). Introducing smoothed taxation within groups of close equals, the AJL decomposition model becomes

$$RE = G_y - G_x^b - G_x^{sw} + G_x^{sw} - G_x^w - R^{AJL},$$

$$RE = V^{AJL} - H^{AJL} - R^{AJL}. \tag{6}$$

In expression (6) the term $V^{AJL} = G_y - (G_x^b + G_x^{sw})$ measures the vertical effect obtained by the difference between the pre-tax Gini coefficient and the Gini coefficient for the income distribution after smoothed tax. $H^{AJL} = G_x^w - G_x^{sw}$ measures horizontal effect given by the difference between the actual within group inequality component and the one which results from a smoothed taxation.

As said above in (6), UL extend AJL model to the case when pre-tax close equals groups are considered instead of exact pre-tax equals; however this adjustment does not solve the problems of measuring both reranking within groups and reranking of entire groups (reranking among group mean incomes). With this aim UL propose a new decomposition method by starting from expression (6)

$$RE = G_y - G_x^{sw} - D_x^b - D_x^w + G_x^{sw} - R^{APK},$$

$$RE = V^{UL} - H^{UL} - R^{APK}, \tag{7}$$

where D_x^w and D_x^b are the within group concentration index and the between group concentration index respectively, $R^{APK} = G_x^b + G_x^w + R^{AJL} - D_x^b - D_x^w$, $V^{UL} = (G_y - G_x^{sw} - D_x^b)$ and $H^{UL} = (D_x^w - G_x^{sw})$. Defining $R^{EG} = G_x^b - D_x^b$ the measure of the reranking among group mean incomes and $R^{WG} = G_x^w - D_x^w$ the measure of the reranking within groups, the Atkinson-Plotnick-Kakwani[4] reranking index R^{APK} may be written as $R^{APK} = R^{EG} + R^{WG} + R^{AJL}$. Each of the three above considered models has some appealing characteristics and some contraindications. In VCL model V^{VCL} measures the vertical effect by comparing pretax and post-tax between groups Gini coefficients, it follows that it does not consider within group inequality existing before and after tax. Moreover the measure the horizontal effect H^{VCL} may fail to capture a within group inequality reduction when moving from pre-tax to post-tax income distribution, as $(G_x^w - G_y^w)$ may result positive even if all post-tax within group Gini coefficients are lower than the corresponding pre-tax ones (see VP, 2007). In AJL

---

[4] Following Kakwani (1984), we observe that, when decomposing by sub-groups each concentration index, we exactly obtain two terms: a within component and a between component.

modified model the horizontal measure H^{AJL} seems to be a proper measure for horizontal effect, as $\left(G_x^w - G_x^{sw} \right)$ is positive when all pre-tax within group Gini coefficients are lower than the corresponding post-tax ones[5] nevertheless this model does not consider properly the reranking effect of taxation. In UL model V^{UL} presents the noticeable advantage that it is a term of a decomposition model which captures all forms of reranking; however once again the interpretation of the horizontal effect is not obvious: as UL notice, H^{UL} may be negative even for relatively small bandwidths, so it should be considered under some arrangements and transformations.

2.1. How the vertical effect should be measured?

When close equals are considered the issue related to the choice of a proper bandwidth arises. As the various components of the decomposition change their relative sizes in accordance with the chosen bandwidth. Thus an arbitrary choice can lead to misleading results (VCL, 2001).

VCL and Kim and Lambert (2008) implicitly assume the vertical effect as a proxy of the potential redistribution of the tax system and they suggest to choose the bandwidth which maximizes V^{VCL} and V^{UL} respectively.

We are not sure that these suggestions always hold. Considering V^{VCL}, it may happen that the bandwidth which maximizes the index is quite large[6], then we can discuss if for that bandwidth the definition of close equals groups holds. In what it concerns V^{UL}, from our experience, this measure may exhibit irregular trend so that[7] the choice of the maximizing bandwidth is not obvious.

VP (2007) suggest to adopt a kind of synthesizing compromise. Following VCL and Kim and Lambert (2008), they assume the vertical effect as a proxy of the potential redistribution and notwithstanding the partial failure of the afore-mentioned indices to measure this effect, they believe that each one may account for it. They observe that V^{AJL} is everywhere lower[8] both than V^{VCL} and V^{UL}[9], and they suggest to choose the bandwidth which satisfies the following criterion:

---

5   This is due to weights associated to pre-tax and post-tax within group Gini indices: they are the same either in G_x^w and in G_x^{sw}, whilst are different in G_y^w (see e.g. Vernizzi and Pellegrino 2007).

6   As we shall see in the pursue, in our data set the maximum V^{VCL} occurs when the bandwidth is 31,000 euro large (Figure 1B).

7   As it is shown in Figure 1A, this measure exhibits an irregular decreasing trend so that, at least in our case, it is not clear how to identify the maximum.

8   VP try to explain why these dominance relationships hold.

9   UL (2008),observe the same relation between the three vertical effect indices.

$$\min\left(\frac{\max\left(\left|V^{VCL}-V^{UL}\right|,\left|V^{VCL}-V^{AJL}\right|,\left|V^{AJL}-V^{UL}\right|\right)}{\min\left(V^{VCL},V^{AJL},V^{UL}\right)}\right) \tag{8}$$

In the pursue we shall deeply analyze how this criterion works in the specific contest of the here considered data set showing that it identifies a bandwidth which is suitable for each of the considered models.

3. Empirical application

In this section we use a micro data set to analyse the relationship among VCL, AJL and UL decompositions considering various bandwidths. To examine how sensitive our decomposition results are to the choice of the bandwidth, we investigate the trend of the measures of vertical, horizontal and reranking effect across three different sets of possible income class widths. Each of the three sets considers one hundred contiguous bandwidths, each bandwidth being a multiple of the minimum in the set. The set are organized as follows with respect to the bandwidth. First set minimum 10 euro, maximum is 1,000 euro. Second set minimum 1000 euro, maximum 100,000 euro. Third set minimum 100,000 euro, maximum 10,000,000 euro. This empirical framework extends our analysis on the complete range of possible bandwidths.

The data derives from AMeRIcA Data Warehouse, which provides demographic and income information for individuals and households resident in the Milanese area. AMeRIcA combines administrative micro data from the tax register of the Milan Revenue Agency with personal data from the Milanese Registry Office. We consider data about Italian Personal Income Tax for the population of individuals resident in the Milanese area. For each taxpayer AMeRIcA contains gross income by source, income tax paid, and the amounts of tax allowances and deductions.

Our attention is focused upon the individual income data collected in 2003. After deleting observations with non-positive gross income, we refer to a population composed by 821260 individuals.

3.1. Results

The values for redistributive effect and reranking effect are not conditioned by the choice of bandwidth. The pre-tax Gini coefficient equals 0.5149659 and the post-tax Gini coefficient is 0.4691416, then the redistributive effect (*RE*) equals 0.0455543. The Atkinson-Plotnick-Kakwani index gives 0.0011725 for the total reranking effect. When we decompose these measures, the magnitudes

of the different decomposition terms depend on the chosen bandwidth. In order to choose a convenient bandwidth, in the following we analyze the behaviour of each of the redistribution effects considering the three proposed decomposition models.

The vertical effect measures

Figure 1A shows that for very small bandwidths V^{UL} dominates V^{VCL} and V^{AJL}, however, enlarging the bandwidth, V^{UL} decreases as V^{VCL} and V^{AJL} increase. For bandwidth larger than 200 euro V^{AJL} becomes distinguishable from V^{VCL} and it slopes down below V^{VCL} line as the bandwidth enlarges. The three measures of vertical effects are close together in correspondence of a bandwidth that approximately ranges from 350 euro to 400 euro; as the bandwidth is widened, V^{VCL} is over V^{UL} line and the distance between V^{VCL} and V^{AJL} increases. For bandwidths larger than 500 euro the trend of the three lines becomes more irregular and V^{AJL} is undistinguishable from V^{UL}.

In Figure 1B V^{UL}, V^{VCL} and V^{AJL} are plotted over a range of bandwidths from 1,000 euro to 100,000 euro. We observe that V^{AJL} and V^{UL} lie on the same line, which is below V^{VCL} line; V^{VCL} continues to increase up to a bandwidth approximately 31,000 euro large, where it amounts to 113.8 percent of RE. V^{VCL} remains greater than RE for bandwidths lower than 58,000 euro, whilst V^{AJL} and V^{UL} becomes lower than RE for bandwidths approximately larger than 11,000 euro. When the bandwidth becomes larger than 100,000 euro, the three measures show a similar decreasing trend.

Figure 1C extends our analysis by considering a very large range for the bandwidth, that varies from 100,000 euro to 10,000,000 euro. We observe that for bandwidths larger than 1,000,000 V^{UL}, V^{VCL} and V^{AJL} capture a negligible percentage of RE; moreover, when the bandwidth tends to its maximum value, the three measures of vertical effect become closer and closer one another and to their zero limit.

Here again we investigate the magnitude of the divergence among V^{UL}, V^{VCL} and V^{AJL} across this complete range of bandwidths by considering the maximum distance among the three computed vertical effect measures. For each bandwidth this maximum distance is defined as follows

$$\Delta_{max} = \max\left(\left|V^{VCL} - V^{UL}\right|, \left|V^{VCL} - V^{AJL}\right|, \left|V^{AJL} - V^{UL}\right|\right) \tag{9}$$

Turning now to the differences between each pair of vertical measures, we observe that $\left(V^{VCL}-V^{AJL}\right)=(G_x^{sw}-G_y^w)$ and $\left(V^{UL}-V^{AJL}\right)=R^{EG}$; so the difference between V^{VCL} and V^{UL} is given by $(G_x^{sw}-G_y^w)$ minus R^{EG}. From Figure 6A we notice that R^{EG} becomes very small for bandwidths larger than 500 euro, then V^{UL} and V^{AJL} become undistinguishable for bandwidths which are 500 euro or more large. Conversely $(G_x^{sw}-G_y^w)$, which is very small for the tiniest bandwidths, increases roughly up to 63,000 euro large bandwidths. As shown earlier (Figures 1A and 1B), V^{VCL} and V^{AJL} lines overlap up to 200 euro large bandwidths, then V^{VCL} line is over V^{AJL} line. For these reasons, when bandwidths are relatively tiny, Δ_{max} is given by R^{EG}, whilst for relatively larger bandwidths it is given by $(G_x^{sw}-G_y^w)$. It results that the descending part of the graph represented in Figure 1A is just R^{EG}, whilst the increasing part is $(G_x^{sw}-G_y^w)$. Figure 2B shows that $(G_x^{sw}-G_y^w)$ presents a reversed U-shape, and in the limit it becomes zero, as shown in Figure 2C.

VP (2007) initially suggest to choose the bandwidth where Δ_{max} is minimum, provided that the maximum among the vertical effect measures is not lower than the lowest among their global maxima over the range of bandwidths; however, after having realized that the second part of the condition does not allow a valid application (see Figure 1A), they modify their criterion suggesting to choose the bandwidth where

$$\Phi = \frac{\max\left(\left|V^{VCL}-V\right|,\left|V^{VCL}-V^{AJL}\right|,\left|V^{AJL}-V\right|\right)}{\min\left(V^{VCL},V^{AJL},V\right)} \tag{10}$$

is minimum.

Using this ratio we relate the magnitude of the difference among the three measures of vertical effect to the measure which captures the potential redistribution less than the others for a given bandwidth. Figures 3A, 3B and 3C report Φ plotted along the entire range of considered bandwidths. Looking at Figure 3A we note that for the starting bandwidth the ratio takes on a value of 0.00073, enlarging the bandwidth, Φ shows a decreasing trend with some irregularities until a bandwidth of 350 euro; for this bandwidth the ratio equals its minimum, then it begins to increase plotting a more regular line than for smaller bandwidths. This rising trend is also confirmed by Figures 3B and 3C which depict an increasing monotonic line for Φ.

We conclude that due to its behaviour, Φ can be considered as a proper indicator to choose a convenient bandwidth. On one side low values of Φ show that V^{UL}, V^{VCL} and V^{AJL} altogether converge to a similar evaluation for the potential redistribution induced by a tax system, and on the other side for low values of Φ we exclude bandwidths which lead to not significant values for V^{UL}, V^{VCL} and V^{AJL}.

We conclude that if the bandwidth 350 is euro large, the three vertical effects give close results for the potential redistributive effect, and these values are higher then the actual redistributive effect.

The horizontal effect measures

The measures of horizontal effect H^{UL}, H^{VCL} and H^{AJL} are plotted over the three bandwidth sets considered above. In Figure 5A H^{VCL} and H^{AJL} show increasing trends: the former up to 31,000 euro large bandwidth, the latter up to 5,000 euro; moreover H^{VCL} remains positive until bandwidths are larger than 57,000 euro, whilst H^{AJL} becomes negative for 12,000 euro large bandwidth. H^{UL} exhibits a different behaviour: it presents positive, even quite low, values, for bandwidths narrower than 150 euro, then H^{UL} oscillates around zero until it starts to decrease till its limit $-RE$. H^{AJL} becomes lower than H^{VCL} for bandwidth larger than 400 Eu. Looking at Figures 5B and 5C, we notice that H^{AJL} remains lower than H^{VCL} as the bandwidth are large, and finally both converge to the common limit $-RE$. In order to investigate the different behaviours of H^{VCL} and H^{AJL}, we recall VP formulae for these measures

$$
\begin{aligned}
H^{VCL} &= G_x^w - G_y^w = \sum_{i=1}^{h}\left(G_{i,x}\frac{n_i^2\mu_i(1-t_i)}{n^2\mu(1-\overline{t})} - G_{i,y}\frac{n_i^2\mu_i}{n^2\mu} \right) \\
H^{AJL} &= G_x^w - G_x^{sw} = \sum_{i=1}^{h}\left(G_{i,x} - G_{i,y} \right)\frac{n_i^2\mu_i(1-t_i)}{n^2\mu(1-\overline{t})}
\end{aligned}
\tag{11}
$$

where μ is mean pre-tax income for the entire population which accounts n income units, μ_i is mean pre-tax income of group i, n_i is the number of income units of group i, $\overline{t}$ is the aggregate tax rate of the entire population, and t_i is the average tax rate of group i. We notice that even if $G_{i,x}$ is lower than $G_{i,y}$, as it is likely to expect assuming a progressive taxation, the corresponding weights may induce the difference between the weighted $G_{i,y}$ and the weighted $G_{i,x}$ to result

positive. Focusing upon these weights, we point out that for a generic group i the pre-tax weight differs from the post-tax one by the ratio $(1-t_i)/(1-\bar{t})$. When $t_i < \bar{t}$ holds, the post-tax weight is greater than the weight before tax and the difference between post-tax weighted Gini coefficient and the pre-tax one may result positive. In a scenario of progressive tax schedule, we assume that lower incomes face an average tax rate lower than the aggregate tax rate $\bar{t}$. Then the greater the number of incomes lower down the distribution, the more numerous the set of taxpayers who face an average tax rate lower than $\bar{t}$. This is the case in presence of skewness, because the income halfway up the distribution is itself below mean income.

From Figure 4, which delineates the density income function for our reference population distribution, we observe that there is evidence of the presence of right skew in the distribution. To explore the effect of the skewness on computing H^{VCL}, we consider Table 1 which reports average tax rates, Gini coefficients, and weighted Gini coefficient both per decile and per cumulative deciles. We define groups of taxpayers using pre-tax deciles to determine group boundaries.

From part (A) of Table 1 we notice that being post-tax Gini coefficient greater than the respective pre-tax one for deciles from the 1th to the 8th, within deciles inequality arises after tax. The inequality rise suggests that relative difference between rich and poor increases within these groups moving from pre-tax to post-tax distribution. This regressive effect derives from horizontal inequity[10] and divergence[11] which occur within the here considered groups. These phenomena relates to tax allowances and deductions adopted by the tax system, which cause departures from the actual tax schedule. It is more like that one of these regressive effect is observed when the departures from the actual tax schedule involve income units with low pre-tax income rather than high pre-tax income by two reasons. First, Italian Personal Income Tax adopts tax allowances which are decreasing with respect to pre-tax income level; then the incidence of tax allowances diminishes as income is increased. Secondly, for a fixed allowance amount, its impact on determining the post-tax income is proportionally greater when considering low pre-tax incomes rather than high ones. We turn our attention on deciles from the 1th to the 8th, and we notice that within each of these groups incomes are low and relatively close; hence it is reasonable that regressive effect occurs within these groups.

---

[10] The concept of horizontal inequity refers to unequal treatment of equals which arises from departures from effective tax schedule.

[11] Divergence occurs where the richer of two individuals obtains a net gain relative to poorer from pre-tax to post-tax distribution.

Table 1. Basic statistic for selected decile groups (pre-tax). Individuals. Milan. 2003.

	INCOME DECILE RANGES										
Decile	1	2	3	4	5	6	7	8	9	10	
Decile upper limit	4826	7401	11197	14307	17136	20301	24419	31017	47280	19057066	
Decile ranges	4826	2575	3796	3110	2829	3165	4118	6598	16263	19009786	
	(A) WITHIN DECILE STATISTICS										Total
G_X	0.408652	0.106611	0.088392	0.060159	0.047637	0.042318	0.037910	0.042961	0.063815	0.346804	/
G_Y	0.403979	0.068549	0.068355	0.040424	0.029914	0.028080	0.030661	0.039578	0.070104	0.376421	/
$G_X - G_Y$	0.004673	0.038062	0.020037	0.019735	0.017723	0.014238	0.007248	0.003384	-0.006290	-0.029620	/
G^W_X	0.000370	0.000278	0.000374	0.000340	0.000325	0.000338	0.000354	0.000482	0.000941	0.012654	0.016545
G^W_Y	0.000313	0.000158	0.000248	0.000201	0.000182	0.000203	0.000265	0.000421	0.001025	0.015462	0.018479
$G^W_X - G^W_Y$	0.000057	0.000120	0.000126	0.000139	0.000143	0.000135	0.000089	0.000061	-0.000085	-0.002810	-0.001934
Average Tax Rate	12.37%	15.19%	12.66%	14.88%	16.21%	17.34%	19.08%	20.97%	24.50%	33.47%	25.10%
Pre-tax Income Share	0.77	2.30	3.62	4.97	6.10	7.25	8.65	10.63	14.63	41.08	100.00
	(B) CUMULATIVE DECILE STATISTICS										
G_X	0.408652	0.335178	0.330968	0.322541	0.312348	0.305688	0.304265	0.311286	0.335181	0.469142	/
G_Y	0.403979	0.324545	0.319302	0.317794	0.312328	0.309234	0.312087	0.323705	0.355384	0.514696	/
$G_X - G_Y$	0.004673	0.010633	0.011667	0.004747	0.000020	-0.003550	-0.007820	-0.012420	-0.020200	-0.045550	/
G^W_X	0.000370	0.002354	0.007683	0.017274	0.031568	0.051743	0.079984	0.121446	0.191593	0.469142	/
G^W_Y	0.000313	0.001996	0.006418	0.014838	0.027754	0.046421	0.073549	0.114710	0.188464	0.514696	/
$G^W_X - G^W_Y$	0.000057	0.000358	0.001265	0.002436	0.003814	0.005322	0.006435	0.006737	0.003129	-0.045550	/
Average Tax Rate	12.37%	14.48%	13.50%	14.09%	14.81%	15.54%	16.45%	17.54%	19.27%	25.10%	/
Pre-tax Income Share	0.77	3.07	6.70	11.67	17.77	25.02	33.67	44.30	58.92	100.00	/

Source: own calculations.

If we focus upon the cumulative analysis, reported in part (B) of the Table 1, we see that the difference between the post-tax and the corresponding pre-tax Gini coefficients becomes negative for the 6th decile, even though the difference calculated for weighted Gini coefficients $\left(G^W_{i,x} - G^W_{i,y}\right)$ remains positive until the 9th decile. As previously observed, this is due the different weights which are applied to $G^W_{9,x}$ and $G^W_{9,y}$; in fact for the 9th decile $t_i < \bar{t}$ holds, then the weight of $G^W_{9,x}$ is greater than $G^W_{9,y}$ one. This explain why H^{AJL} becomes negative much before H^{VCL}, being for the former the pre-tax and the post-tax Gini coefficients weighted by the same coefficients within each group, which is not the case for the latter.

When considering cumulative deciles groups, the number and the disparities of incomes allocated to each group increase, then the incidence of the departures from effective tax schedule diminishes. Looking at the cumulative statistics in Table 1, part B, we observe that the within group inequality starts to decrease from the 6th decile, however the weighted post-tax Gini coefficients remain greater than the corresponding pre-tax ones until the 9th decile. This is due to the effect of the average tax rate which remains lower than the aggregate tax rate until the 9th decile included. Moreover we observe that the magnitude of the positive difference between the weighted post-tax Gini coefficient and the pre-tax one increases together with the population share included within a same group, and its maximum is for the 8th cumulative decile; conversely this difference decreases when considering the 9th cumulative decile, and it becomes negative for the 10th cumulative decile. Our results indicate that H^{VCL} takes on its maximum value for the bandwidth which leads to include within a same group the lower 80 percent incomes, which happens when the bandwidth is a bit larger than 31,000 euro; when the bandwidth is further enlarged, H^{VCL} starts to decrease and it becomes negative for bandwidths larger than 58,000 euro. This behaviour can be explained by two reasons. First, when the bandwidth is larger than 31,000 euro lower 80 percent incomes are confounded with a part of top 20 percent income distribution within a same group and, consequently, both disparities and average tax rate increase by including additional incomes from the top 20 percent of income distribution. In addition, according to the progressivity of tax schedule, liabilities are proportionally higher for these incomes, hence the magnitude of inequality reduction within the group should increase when higher incomes are added to the group. Second, enlarging the bandwidth, it is likely to expect that the number of income units included in any group increases, and on the other hand the number of identified groups diminishes. Then the within component of Gini coefficient better captures the disparities among top 10 percent incomes which are widespread along the right tail of the income distribution.

We explain the behaviour of H^{VCL} and H^{AJL} focusing on the proportions of population of taxpayers which are split into sub-groups when defining the bandwidth; however our main goal is to separate the redistributive effect by selecting a convenient bandwidth. When considering the selected bandwidth (350 euro), H is close to zero, whereas H^{AJL} results positive; then we suggest to consider H^{AJL} as horizontal inequity measure.

The reranking effect measures

Figure 6 outlines the behaviour of R^{APK} decomposition across the three sets of bandwidths considered here. Looking at Figure 6A, we observe that for nar-

row bandwidths R^{AJL} represents a high percentage of R^{APK}, at the starting bandwidth R^{AJL} equals to the 96.6 percent of R^{APK} and it reaches its maximum value (97.1 percent) if the bandwidth is 20 euro large, then R^{AJL} decreases as the bandwidth is widened. While R^{WG} exhibits an upward trend which seems to be approximately proportional to the bandwidth, R^{EG} quickly falls by enlarging the bandwidth. If the bandwidth equals 10 euro, R^{EG} percentage of R^{APK} is 2.89 percent and it is higher than R^{WG} one, when the bandwidth is 300 euro or more large, R^{EG} is lower than 0.5 percent of R^{APK}. From Figure 6B we observe that, as R^{AJL} decrease of a certain percentage of R^{APK}, R^{WG} raises of the same percentage amount. This results from the behaviour of R^{EG} whose percentage of R^{APK} is close to zero (as noted earlier, looking at Figure 6A), for this reason R^{EG} is not delineated in Figure 6B. Figure 6C shows that R^{WG} approximates the value of R^{APK} when the bandwidth is 800,000 euro or more large.

If we consider a bandwidth 350 euro large, we obtain only two reranking components, that are R^{WG} and R^{AJL}, because for this bandwidth R^{EG} represents a negligible percentage of R^{APK}.

4. Conclusion

The purpose of this work is to examine the decomposition of the redistributive into vertical, horizontal and reranking components in order to select the bandwidth that is used to split the taxpayers resident in Milan into sub-groups. The choice of the bandwidth determines the magnitude of these components and the relationship among them. In this paper we apply three alternative methods to decompose the redistributive effect obtained moving from pre-tax to post-tax income distribution: the model suggested by Aronson, Johonson and Lambert (1994) and refined by Urban and Lambert (2008), the method proposed by van de Ven, Creedy and Lambert (2001), and the one recently introduced in Urban and Lambert. We recall Vernizzi and Pellegrino (2007) methodology for selecting the bandwidth in order to check whether their criterion identify a proper bandwidth to decompose the redistributive effect which occurs with respect to Milan micro data.

Our findings confirm that the criterion is adequate to set a proper bandwidth to decompose the redistributive effect. From the comparison of the differences among V^{VCL}, V^{AJL}, V^{UL} across a broad set of possible bandwidths, we notice that the criterion suggests to choose a bandwidth equal to 350 euro, which ensures

that these alternative measures of vertical effect give coherent outcomes and capture a potential redistribution.

The decomposition of the reranking effect indicates that the reranking which involves group mean incomes (entire-group reranking) becomes a negligible percentage of the total reranking effect when the bandwidths are 300 euro or more large. Then, for larger intervals, any rise of the within-group reranking term is balanced by a fall of the same amount in reranking component that accounts for reranking between income units belonging to different groups.

We also obtain results to examine how the skewness of the income distribution affects the determination of the horizontal effect. We explain the behaviour of H^{VCL} and H^{AJL} focusing on the proportions of population of taxpayers which are split in sub-groups when defining the bandwidth. Our analysis shows that the weights attributed to sub-groups are shifted from pre-tax to post-tax distribution for effect of the sub-groups average tax rates. This determines that the post-tax within Gini coefficient is greater than both the pre-tax within Gini coefficient and smoothed within Gini coefficient, even if sub-groups post-tax Gini coefficients are lower than the corresponding pre-tax ones. We find evidence that lower tax rates are associated with lower 80 percent incomes of the distribution, and hence higher values of H^{VCL} are obtained when lower 80 percent incomes of the distribution are included within the same group.

Figure 1: Vertical effect plotted over different ranges: (A) from 10 euro to 1,000 euro – increasing step 10 euro; (B) from 1,000 euro to 100,000 euro – increasing step 1,000 euro; (C) from 100,000 euro to 10,000,000 euro – increasing step 100,000 euro.
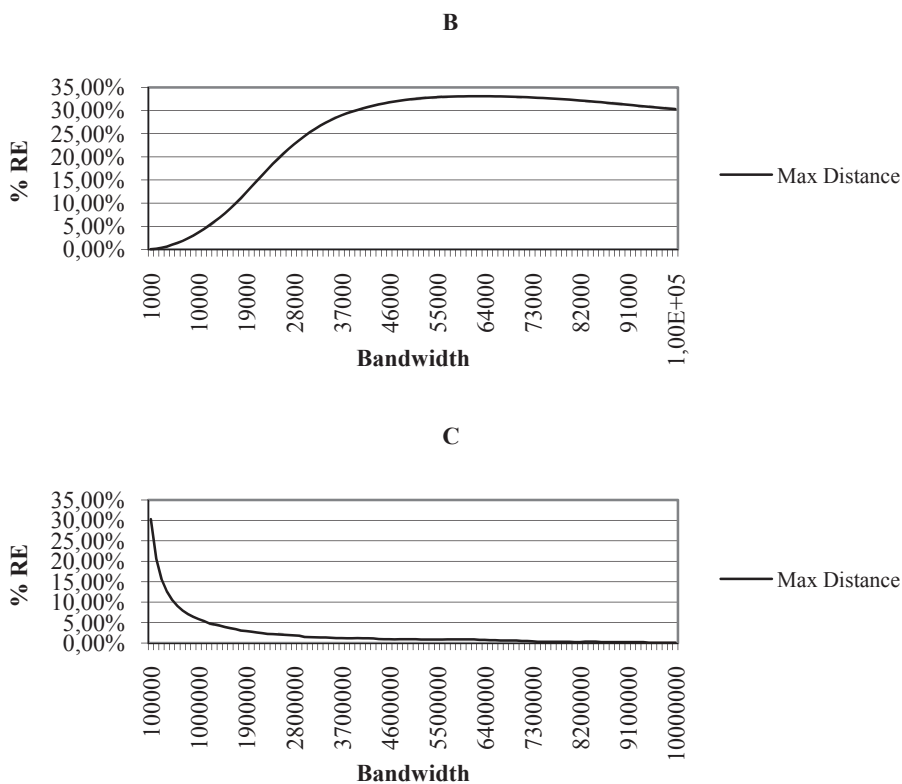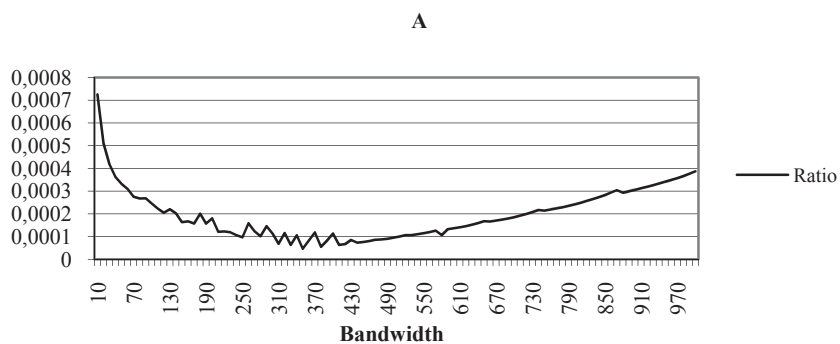
A

B

C
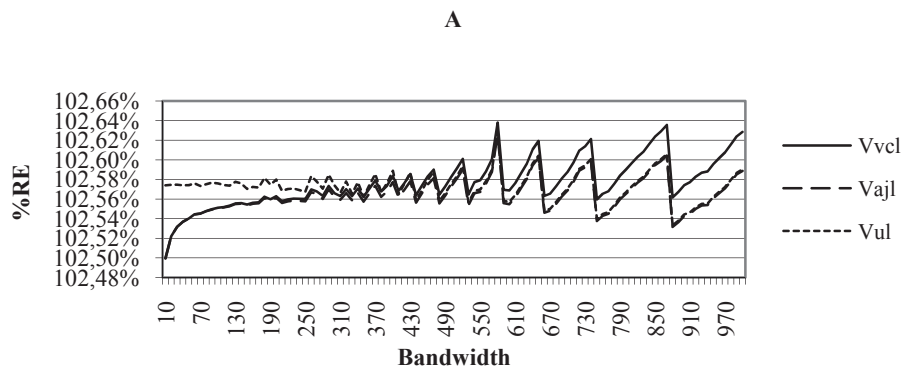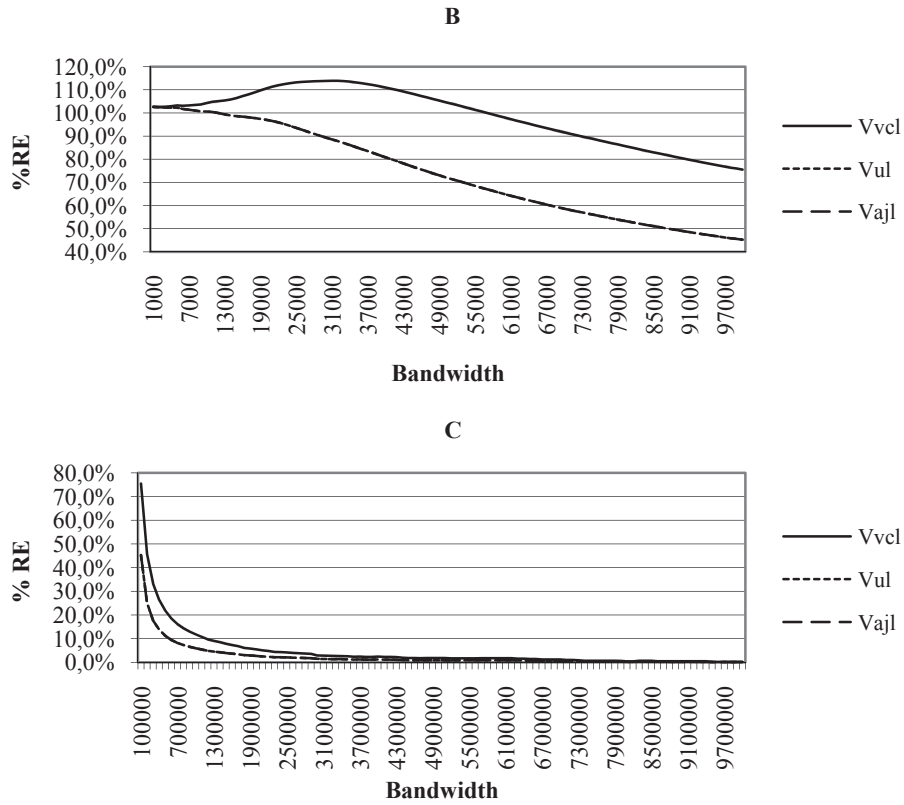
Figure 2: Maximum distance plotted over different ranges: (A) from 10 euro to 1,000 euro – increasing step 10 euro; (B) from 1,000 euro to 100,000 euro – increasing step 1,000 euro; (C) from 100,000 euro to 10,000,000 euro – increasing step 100,000 euro.
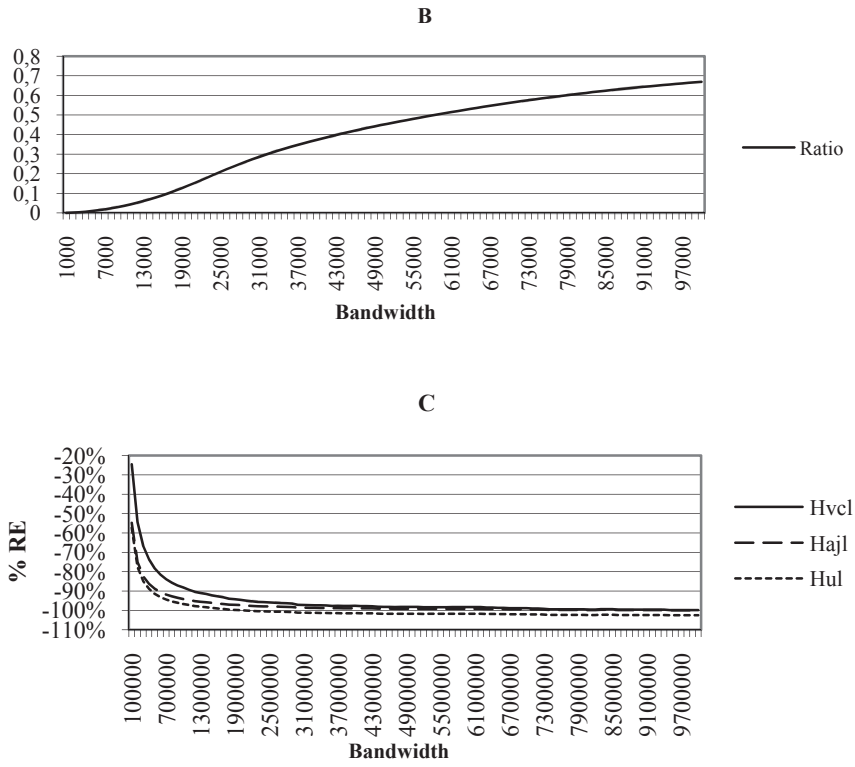
A

B

C
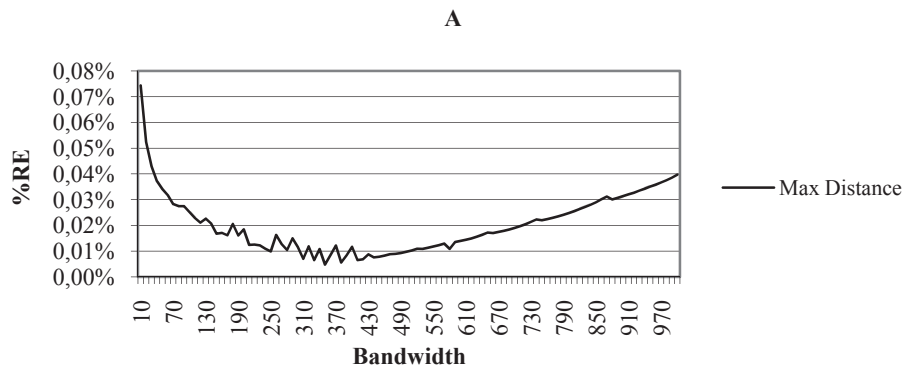
Figure 3: Ratio plotted over different ranges: (A) from 10 euro to 1,000 euro – increasing step 10 euro; (B) from 1,000 euro to 100,000 euro – increasing step 1,000 euro; (C) from 100,000 euro to 10,000,000 euro – increasing step 100,000 euro.
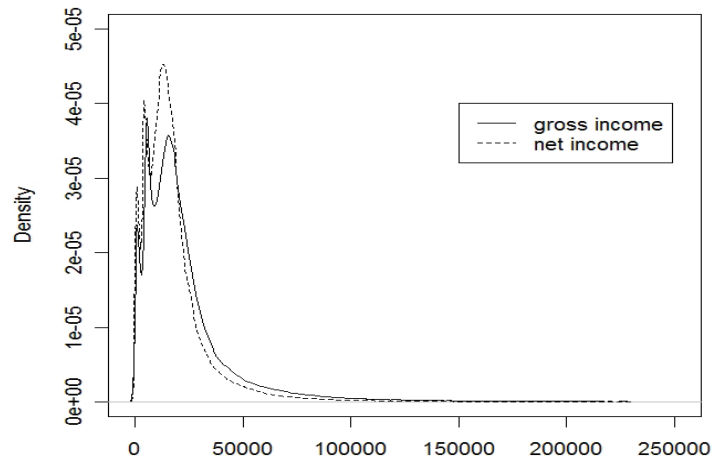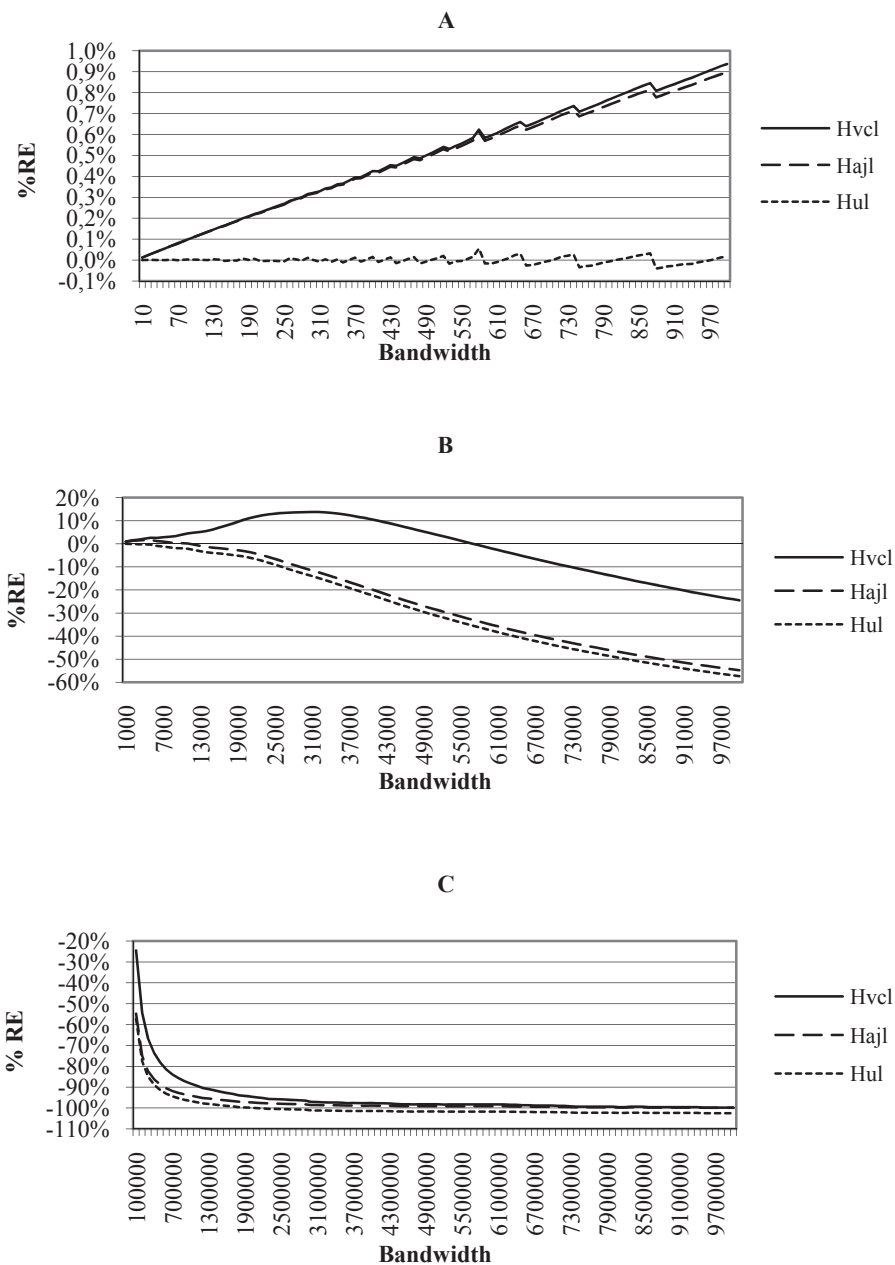
A

B

C

Figure 4: Pre-tax income density function and post-tax income density function. Individuals. 2003.

Figure 5: Horizontal effect plotted over different ranges: (A) from 10 euro to 1,000 euro – increasing step 10 euro; (B) from 1,000 euro to 100,000 euro – increasing step 1,000 euro; (C) from 100,000 euro to 10,000,000 euro – increasing step 100,000 euro.

Figure 6: Reranking effect plotted over different ranges: (A) from 10 euro to 1,000 euro – increasing step 10 euro; (B) from 1,000 euro to 100,000 euro – increasing step 1,000 euro; (C) from 100,000 euro to 10,000,000 euro – increasing step 100,000 euro.
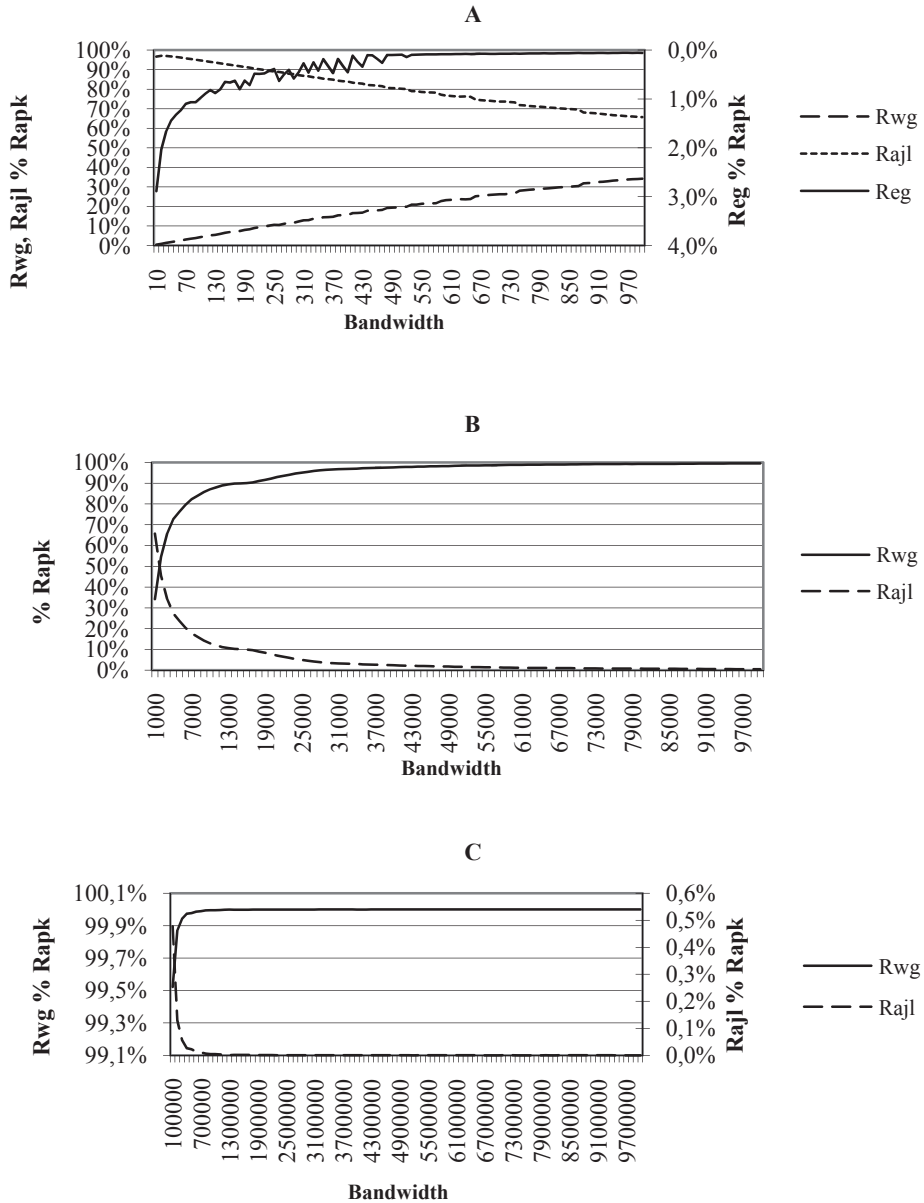
Literature

[1] Aronson R.J., Lambert P.J., (1993), "Inequality decomposition analysis and the Gini coefficient revisited", *The Economic Journal*, 103, pp. 1221–1227.

[2] Aronson R.J., Lambert. P.J., (1994), "Decomposing the Gini coefficient to reveal the vertical, horizontal and reranking effects of income taxation", *National Tax Journal*, 47, pp. 273–294.

[3] Aronson R.J., Johnson P.J., Lambert P.J., (1994), "Redistributive effect and unequal income tax treatment", *The Economic Journal*, 104, pp. 262–270.

[4] Atkinson A. B., (1980), "Horizontal equity and the distribution of the tax burden", in *The economics of taxation* Henry J.A. and Boskins M.J. (Eds.), 3–18, Washington, DC: Brookings Institution.

[5] Dagum C. (1997), "A new approach to the decomposition of Gini income inequality ratio", *Empirical Economics*, 22, pp. 515–531.

[6] Kakwani N. (1984) "On the measurement of tax progressivity and redistributive effect of taxes with application to horizontal and vertical equity", *Advances in Econometrics*, 3, 149–168.

[7] Kim K., Lambert P.J., (2008), "Redistributive Effect of U.S. Taxes and Pubblic Transfers, 1994–2004", *Public Finance Review*, 20, n. 10, pp.1–24.

[8] Plotnick R., (1981), "A measure of horizontal inequity", *Review of Economics and Statistics*, 63, pp.283–287.

[9] Urban I., Lambert P.J., (2008), "Redistribution, horizontal inequity and reranking: how to measure them properly", *Public Finance Review*, 36, n. 5, pp. 563–587.

[10] van de Ven J., Creedy J., Lambert P.J., (2001), "Close equals and calculation of the vertical, horizontal and reranking effects of taxation", *Oxford Bulletin of Economics and Statistics*, 63, 3, pp. 381–394.

[11] Vernizzi A., Pellegrino S., (2007), "On determining "close equals groups" in decomposing redistributive and reranking effects", *SIEP*, WP 2007–602. Available at http://ideas.repec.org/p/bep/unimip/1075.html.

Streszczenie

Wybór rozpiętości dochodów w dekompozycji efektu redystrybucji na podstawie analizy danych pochodzących z AMeRIcA

Wykorzystując dekompozycję efektu redystrybucji dochodów na skutek opodatkowania, należy zdefiniować grupy jednostek charakteryzujących się podobnymi dochodami. W artykule zostały przedstawione sposoby rozwiązania tego problemu proponowane przez różnych autorów. Następnie na podstawie rzeczywistych danych przeprowadzono empiryczną analizę porównawczą proponowanych rozwiązań.