Duclos-Jalbert-Araar decomposition of redistributive effect: implementation issues

Ivica, Urban

Institut za javne financije

11 January 2011

Duclos-Jalbert-Araar decomposition of redistributive effect: implementation issues

IVICA URBAN

Institut za javne financije, Zagreb, Croatia

Version: January 11, 2011; **Comments invited** (ivica.urban@ijf.hr)

Abstract

When there is a large number of pre-fiscal income equals in the sample, the redistributive effects based on the Gini index may be inappropriately estimated. Due to its roots in the Gini framework, the problem also appears in the application of the Duclos, Jalbert and Araar (2003) (DJA) decomposition of redistributive effects into vertical, classical horizontal inequity and reranking effects. This paper explains how the DJA implementation procedure must be adapted to produce correct estimates of different effects. The procedure is first illustrated on the 12-units hypothetical population and then applied to real data for the Croatian fiscal subsystem comprised of social security contributions, personal income tax, public pensions and cash social benefits.

Keywords: redistributive effect, vertical equity, horizontal equity, pre-fiscal equals

JEL: D63, H22, H23

1. Introduction

Duclos, Jalbert and Araar (2003) (DJA) have designed a comprehensive model for the measurement of redistributive effect (RE).[1] The DJA model decomposes RE into vertical, classical horizontal inequity (CHI) and reranking effects. It is built into the framework of the Atkinson-Gini social welfare functions, which makes it different from its older competitor, the widely acknowledged Aronson, Johnson and Lambert's (1994) (AJL) decomposition of RE, which is set up in the Gini environment. The Atkinson-Gini social welfare function first converts

[1] RE is a change in inequality occurring in a transition from pre-fiscal to post-fiscal income. Generally, a pre-fiscal (post-fiscal) income is the income before (after) taxes and benefits.

incomes into utilities employing the Atkinson (1970) utility function, and then aggregates them using rank-dependent weights, which underlie the extended Gini coefficients proposed by Yitzhaki (1979).

Despite its great measurement potential, the DJA model has not yet become widely employed by practitioners.[2] Among the reasons for such condition are implementation difficulties. The model asks for the estimation of expected post-fiscal incomes (EPI) at different points of the pre-fiscal income distribution (PRFID), which requires certain statistical expertise related to data smoothing and curve fitting methods. Urban (2010) explains why inaccurate estimates of EPI lead to misleading values of indicators in the DJA model. A simple test is suggested to check the appropriateness of EPI estimates. Specifically, for certain combinations of two ethical parameters of the Atkinson-Gini framework, the CHI effect should collapse to zero; if this is not the case, the practitioner should re-evaluate the EPI curve.

This paper reveals another challenge for practitioners applying the DJA model: if there is a large number of pre-fiscal income equals in the sample, common procedures for index computation will lead to flawed results.

Researchers in the field of income redistribution know that exact pre-fiscal income equals are rare in the data samples. Still, there is one important exception: in the analysis of tax-and-benefit systems, it may occur that many income units receive only income from government in the form of benefits, which implies that their pre-fiscal income is zero. Such an example came into sight in the study of Croatian individual taxes and cash social benefits, when one of the scenarios treated public pensions as social benefits. The consequence was that more than one tenth of all income units in the sample had zero pre-fiscal income.

[2] Bilger (2008) is the only published research using the DJA model.

A presence of a large number of pre-fiscal equals concentrated in one point of the PRFID can affect the estimates of various redistributive effects unless the appropriate adaptations are made. Urban (2010) explains how the post-fiscal income vector and pre-fiscal income quantiles must be adjusted to enable the correct estimation of the concentration coefficient of post-fiscal income. Otherwise, both the vertical and horizontal inequity terms in the Kakwani's (1984) (K84) decomposition of RE, also set in the Gini environment, would be underestimated.

Since K84 and DJA share the same Gini environment, in which incomes (utilities) are aggregated using rank-dependent weights, the problem of correct estimation in the presence of pre-fiscal equals also emerges in the DJA model. In the latter model, a failure to make adaptations can lead to a wrong assessment of the relative importance of CHI versus reranking. Furthermore, it can mistakenly convince the analyst that the estimate of EPI is inaccurate.

This paper carefully explains the data manipulation and calculation procedures needed to obtain the indices of the DJA model in the case when there is a large number of (zero or non-zero) pre-fiscal equals is in the sample. However, it can be useful to all practitioners employing standard methodologies in the field of income redistribution. Together with Urban (2010), it can serve as a 'manual' for practitioners applying the DJA methodology together with the original DJA's work.

Another peculiarity has been observed during the study of the Croatian fiscal system. Specifically, the estimated EPI curve is not increasing in a pre-fiscal income across the whole distribution of pre-fiscal incomes. This implies that the counterfactual fiscal system defined by EPI does not fully eliminate HI; a certain amount is left in the form of reranking. The implications are also discussed below.

Section 2 describes the procedures of data preparation and the calculation of various elements of the DJA model (counterfactual incomes, utilities, weights, welfare and inequality indices), ending with a short exposition of the DJA model. Section 3 applies the procedures to two different examples: a hypothetical one with a population of 12 units, and the other one, using real data on the Croatian system of personal taxes and social benefits in 2008. Different components obtained for real data are further analyzed and interpreted. Section 4 concludes.

2. Calculation of indices

2.1. Basic data preparation

A typical research uses the following data for a household or family k: (a) unequivalized pre- and post-fiscal incomes, $\dot{X}_k$ and $\dot{N}_k$, (b) survey frequency weights f_k, and (c) equivalence factors β_k.[3] Equivalized (henceforth, *equivalized* is omitted) pre- and post-fiscal incomes are $X_k = \dot{X}_k / \beta_k$ and $N_k = \dot{N}_k / \beta_k$, and frequency weights are $\phi_k = f_k \beta_k$. In the following analysis, we consider that household k has β_k 'equivalent' members instead of some number u_k of 'real' individuals; thus, each equivalized income pair (X_k, N_k) will be counted ϕ_k and not $f_k u_k$ times.

The matrix $M_k^0 = \{X_k, N_k, \phi_k\}$ is sorted lexicographically, first, in increasing order of pre-fiscal income and then, within each group of pre-fiscal equals, in increasing order of post-fiscal income. Thus, we obtain the matrix $M_i^x = \{X_i^x, N_i^x, \phi_i^x\}$. Sorting M_k^0 in increasing order of post-fiscal income, $M_i^n = \{X_i^n, N_i^n, \phi_i^n\}$ is obtained. For the implementation of the DJA model we

[3] For an explanation of these items, see the concrete example of Croatian data in Section 3.2.

extract from M_k^x and M_k^n the income vectors X_i^x, N_i^x, N_i^n, and the frequency weights ϕ_i^x and ϕ_i^n.

Another vector is needed for the DJA model, the one that contains the estimates of EPI for each value of pre-fiscal income X_i^x. We denote this counterfactual post-fiscal income vector as N_i^E. To obtain it, we must smooth a dataset (X_i^x, N_i^x), i.e. approximate the mean response curve m^E in the regression relationship $N_i^x = m(X_i^x) + \delta_i$. The basic form of the curve m is chosen by the analyst from a great variety of possible choices, such as OLS polynomial regressions, kernel regressions, local polynomial regressions, Gini regressions, Fourier transformations, etc. Let $\hat{m}^E(X_i^x)$ be the correct approximation of m^E; then, $N_i^E = \hat{m}^E(X_i^x)$.

2.2. Utilities

Incomes are converted into utilities according to the Atkinson (1970) utility function with the inequality aversion parameter ε. When $\varepsilon \neq 1$, utilities of incomes Y_i^z are:[4]

$$U(Y_i^z, \varepsilon) = \frac{(Y_i^z)^{1-\varepsilon}}{1-\varepsilon} \tag{1}$$

Analogously to (1), utilities $U(X_i^x, \varepsilon)$, $U(N_i^x, \varepsilon)$, $U(N_i^E, \varepsilon)$ and $U(N_i^n, \varepsilon)$ are obtained.

Now, for each value ε we can estimate the regression relationship $U(N_i^x, \varepsilon) = m^P(X_i^x) + \delta_i$ to obtain the approximation $\hat{m}^P(X_i^x)$ and another vector of fitted values $U_{i,\varepsilon}^P = \hat{m}^P(X_i^x)$. It shows the expected post-fiscal utilities at different points of the PRFID.

[4] We omit the specific formula that must be used when $\varepsilon = 1$ to keep the exposition more simple.

If N_i^E is properly estimated, we should have that:

$$\pi\sum_i^s \varphi_i^x N_i^E = \pi\sum_i^s \varphi_i^x N_i^x = \pi\sum_i^s \varphi_i^n N_i^n \quad (2)$$

where $\pi = (S)^{-1}$. Equation (2) indicates that the means of all post-fiscal income vectors are equal. Similarly, the following average utilities of post-fiscal incomes are also equal if $U_{i,\varepsilon}^P$ is properly estimated:

$$\pi\sum_i^s \varphi_i^x U_{i,\varepsilon}^P = \pi\sum_i^s \varphi_i^x U(N_i^E,\varepsilon) = \pi\sum_i^s \varphi_i^x U(N_i^x,\varepsilon) = \pi\sum_i^s \varphi_i^n U(N_i^n,\varepsilon) \quad (3)$$

2.3. Quantile estimates

We turn to the estimation of the quantiles of PRFID, $\hat{p}_i^x$, and the weights needed for the computation of welfare indices, $\hat{\varpi}_i^{x,\nu}$. They can be obtained in the following way:

$$\hat{p}_i^x = (2S)^{-1}\sum_{j=1}^i (\varphi_j^x + \varphi_{j-1}^x) \quad (4)$$

$$\hat{\varpi}_i^{x,\nu} = (S)^{-1}\nu(1-\hat{p}_i^x)^{\nu-1} \quad (5)$$

where $\phi_0^x = 0$, $S = \sum_{j=1}^s \phi_j^x$, s is the number of sample units and ν is an ethical parameter determining the weights of the Gini social welfare function. Analogously, $\hat{p}_k^n$ and $\hat{\varpi}_k^{n,\nu}$ are obtained.

Alternatively, when $\phi_i^x = 1$ for all i, we can rather use the following formula for the weights $\hat{\varpi}_i^{x,\nu}$:

$$\hat{\varpi}_i^{x,\nu} = s^{-\nu}[(s-i+1)^\nu - (s-i)^\nu] \quad (6)$$

Now, assume that $X_i^x = 0$ for $i = 1,\ldots,q$; these are our q exact equals with zero pre-fiscal income. Observe that the weights $\hat{\varpi}_i^{x,\nu}$ are strictly decreasing in $\hat{p}_i^x$; it means that $\hat{\varpi}_1^{x,\nu} > \ldots > \hat{\varpi}_q^{x,\nu}$, i.e. pre-fiscal equals obtain different weights. However, this contradicts the very notion of the indices based on rank-dependent weights: if there are units with the same pre-fiscal income, they belong to the same quantile, and equal weights should be ascribed to them. Therefore, we must obtain a new set of weights, $\widehat{\varpi}_i^{x,\nu}$:

$$\widehat{\varpi}_i^{x,\nu} = \begin{cases} (\sum_{j=1}^{q} \varphi_j^x)^{-1} \sum_{c=1}^{q} \varphi_c^x \cdot \hat{\varpi}_c^{x,\nu} & \text{for } i = 1,\ldots,q \\ \hat{\varpi}_i^{x,\nu} & \text{for } i > q \end{cases} \tag{7}$$

Of course, the analogous procedure should be applied across the whole PRFID to account for all groups of pre-fiscal equals revealed by the inspection of X_i^x.

2.4. "Welfare" indices

The estimate of the Gini-Atkinson welfare index for incomes Y_i^z and weights $\hat{\varpi}_k^{x,\nu}$ is the following:

$$\hat{W}(Y_i^z, \varepsilon, \nu; \hat{\varpi}_j^{z,\nu}) = \sum_{j=1}^{s} U(Y_i^z, \varepsilon) \cdot \varphi_j^z \cdot \hat{\varpi}_j^{z,\nu} \tag{8}$$

Combining different utilities and sets of weights, we obtain the welfare index $\hat{W}(N_i^n, \varepsilon, \nu; \hat{\varpi}_i^{n,\nu})$, and all other indices between which the following relationships exist:

$$\hat{W}(X_i^x, \varepsilon, \nu; \hat{\varpi}_i^{x,\nu}) = \widehat{W}(X_i^x, \varepsilon, \nu; \widehat{\varpi}_i^{x,\nu}) \tag{9}$$

$$\hat{W}(N_i^E, \varepsilon, \nu; \hat{\varpi}_i^{x,\nu}) = \widehat{W}(N_i^E, \varepsilon, \nu; \widehat{\varpi}_i^{x,\nu}) \tag{10}$$

$$\hat{W}(N_i^x, \varepsilon, \nu; \hat{\varpi}_i^{x,\nu}) < \widehat{W}(N_i^x, \varepsilon, \nu; \widehat{\varpi}_i^{x,\nu}) \tag{11}$$

$$\widehat{W}(U_{i,\varepsilon}^{P},\varepsilon,\nu;\widehat{\varpi}_i^{x,\nu}) = \hat{W}(U_{i,\varepsilon}^{P},\varepsilon,\nu;\hat{\varpi}_i^{x,\nu}) = \widehat{W}(N_i^{x},\varepsilon,\nu;\widehat{\varpi}_i^{x,\nu}) \tag{12}$$

Identities (9), (10) and (12) tell us that for vectors X_i^x, N_i^E and $U_{i,\varepsilon}^P$ it does not matter whether $\hat{\varpi}_i^{x,\nu}$ or $\widehat{\varpi}_i^{x,\nu}$ is used, simply because these vectors have equal values for $i = 1,\ldots,q$.

The situation is different for N_i^x; its welfare index for $\hat{\varpi}_i^{x,\nu}$ is lower than when $\widehat{\varpi}_i^{x,\nu}$ are used. In fact, the whole procedure of weights adaptation is introduced in order to properly estimate the welfare index for N_i^x. Why does the index $\hat{W}(N_i^x,\varepsilon,\nu;\hat{\varpi}_i^{x,\nu})$ underestimate the true welfare? To answer this question, we have to recall the meaning of the weights in the model. They assign to each pre-fiscal quantile p relative importance to the overall welfare of incomes situated at this quantile. If $X_i^x = 0$ for $i = 1,\ldots,q$, then all these units' incomes should receive equal importance, i.e. equal weight. For the weights $\hat{\varpi}_i^{x,\nu}$ this is not the case, because units with i closer to q will gain lower weights than those closer to 1, and since incomes N_i^x within each group of exact equals are sorted in ascending order, the result (11) follows.[5]

Somewhat less obvious seems to be the identity (11), saying that $U_{i,\varepsilon}^P$ and N_i^x (if proper weights are used) result in equal welfare indices. Recall that $U_{i,\varepsilon}^P$ averages the utilities $U(N_i^x,\varepsilon)$ at some pre-fiscal quantile p; therefore, the sum of $U_{i,\varepsilon}^P$ for all income units at p should be equal to the sum of $U(N_i^x,\varepsilon)$.

---

[5] A simple example can demonstrate this algebraic effect. Three incomes are 10, 20 and 40; their respective weights A are 3, 2 and 1, while the weights B are 2, 2 and 2 respectively. The sum-product of incomes and weights A (weights B) is 110 (120).

2.5. "Inequality" indices

The inequality index $\hat{I}(Y_i^z,\varepsilon,\nu;\hat{\varpi}_j^{z,\nu})$ is calculated as follows:

$$\hat{I}(Y_i^z,\varepsilon,\nu;\hat{\varpi}_j^{z,\nu}) = 1-[(1-\varepsilon)\hat{W}(Y_i^z,\varepsilon,\nu;\hat{\varpi}_j^{z,\nu})]^{\frac{1}{1-\varepsilon}} / \hat{\mu}^Y \tag{13}$$

where $\hat{\mu}^Y$ is the mean of the income vector Y_i^z. For convenience reasons, we will use the abbreviated version $\hat{I}(Y_i^z)$ or $\overset{\frown}{I}(Y_i^z)$, depending on whether the weights $\hat{\varpi}_j^{z,\nu}$ or $\overset{\frown}{\varpi}_j^{z,\nu}$ are used. Following (13), by using welfare $\hat{W}(X_i^x,\varepsilon,\nu;\hat{\varpi}_i^{x,\nu})$ and the mean pre-fiscal income $\hat{\mu}^X$, we obtain inequality indices of pre-fiscal income $\hat{I}(X_i^x)=\overset{\frown}{I}(X_i^x)$. Similarly, inequality indices $\hat{I}(N_i^n;\hat{\varpi}_i^{n,\nu})$, $\hat{I}(N_i^E)=\overset{\frown}{I}(N_i^E)$, $\hat{I}(U_{i,\varepsilon}^P)=\overset{\frown}{I}(U_{i,\varepsilon}^P)$, $\hat{I}(N_i^x)$ and $\overset{\frown}{I}(N_i^x)$ are obtained by using corresponding welfare indices, weights and the mean post-fiscal income $\hat{\mu}^N$.

The previous section discussed the identity of welfare indices calculated for N_i^x and $U_{i,\varepsilon}^P$. It follows that the corresponding welfare indices are also equal: $\overset{\frown}{I}(N_i^x)=\hat{I}(U_{i,\varepsilon}^P)$. This property has a practical implication that we do not need to estimate $U_{i,\varepsilon}^P$ and $\hat{I}(U_{i,\varepsilon}^P)$ at all; instead, we calculate $\overset{\frown}{I}(N_i^x)$.

2.6. DJA decomposition

The difference $\hat{\Delta}(\varepsilon,\nu)=\hat{I}(X_i^x)-\hat{I}(N_i^n)$ is the basic measure of inequality change caused by a fiscal system – the redistributive effect. The DJA model decomposes RE into three terms, as shown by the following equation, which can be written in two ways, given that $\overset{\frown}{I}(N_i^x)=\hat{I}(U_{i,\varepsilon}^P)$:

$$\begin{aligned}\hat{\Delta}(\varepsilon,\nu) &= [\hat{I}(X_i^x)-\hat{I}(N_i^E)]-[\overset{\frown}{I}(N_i^x)-\hat{I}(N_i^E)]-[\hat{I}(N_i^n)-\overset{\frown}{I}(N_i^x)] = \\ &= [\hat{I}(X_i^x)-\hat{I}(N_i^E)]-[\hat{I}(U_{i,\varepsilon}^P)-\hat{I}(N_i^E)]-[\hat{I}(N_i^n)-\hat{I}(U_{i,\varepsilon}^P)]\end{aligned} \tag{14}$$

The three differences in brackets, $\hat{V}^{DJA}(\varepsilon,\nu)=\hat{I}(X_i^x)-\hat{I}(N_i^E)$, $\hat{C}^{DJA}(\varepsilon,\nu)=\widehat{I}(N_i^x)-$ $-\hat{I}(N_i^E)=\hat{I}(U_{i,\varepsilon}^P)-\hat{I}(N_i^E)$ and $\hat{R}^{DJA}(\varepsilon,\nu)=\hat{I}(N_i^n)-\widehat{I}(N_i^x)=\hat{I}(N_i^n)-\hat{I}(U_{i,\varepsilon}^P)$, are the vertical, CHI and reranking effects of the DJA model.

We can divide $\hat{R}^{DJA}(\varepsilon,\nu)$ into two parts:

$$\hat{R}^{DJA}(\varepsilon,\nu)=[\hat{I}(N_i^n)-\hat{I}(N_i^x)]+[\hat{I}(N_i^x)-\widehat{I}(N_i^x)] \tag{15}$$

The difference $\hat{I}(N_i^x)-\widehat{I}(N_i^x)$ is positive and shows by how much the DJA reranking effect is underestimated if $\hat{I}(N_i^x)$ is employed instead of $\widehat{I}(N_i^x)$. At the same time, and by the same amount, the DJA CHI effect will be overestimated. The difference $\hat{R}^{und}(\varepsilon,\nu)=\hat{I}(N_i^x)-\widehat{I}(N_i^x)$ can be called the 'underestimation of the reranking term'. Its magnitude will be revealed in the following hypothetical and real examples. The other difference, $\hat{R}^{res}(\varepsilon,\nu)=\hat{I}(N_i^n)-\hat{I}(N_i^x)$, can simply be referred to as the 'residual of the reranking term'.

3. Two examples

3.1. Hypothetical data example

The first example employs a hypothetical population of twelve income units. Table 1 presents most of the vectors needed for the computation of different indices. The first six units are pre-fiscal exact equals with zero pre-fiscal income, while the remaining units have different pre-fiscal incomes X_i^x. The two sets of weights are presented; first, the original ones, $\hat{\varpi}_i^{x,\nu}$, obtained by using (6), which assume that all the units have different pre-fiscal incomes; the weights $\widehat{\varpi}_i^{x,\nu}$

are the same as the original ones, except for the units #1 to #6, which are equal to $\widehat{\varpi}_i^{x,\nu} = \sum_{c=1}^{6} \hat{\varpi}_c^{x,\nu} / 6$, as the rule (7) demands.

TABLE 1
HYPOTHETICAL EXAMPLE (UTILITIES FOR $\varepsilon = 0.5$; WEIGHTS FOR $\nu = 2$)

#	$\hat{\varpi}_i^{x,\nu}$	$\widehat{\varpi}_i^{x,\nu}$	X_i^x	N_i^x	N_i^n	N_i^E	$U(X_i^x,\varepsilon)$	$U(N_i^x,\varepsilon)$	$U(N_i^n,\varepsilon)$	$U(N_i^E,\varepsilon)$	$U_{i,\varepsilon}^P$
1	0.160	0.125	0	10	10	52	0.00	6.32	6.32	14.43	13.70
2	0.146	0.125	0	20	20	52	0.00	8.94	8.94	14.43	13.70
3	0.132	0.125	0	30	30	52	0.00	10.95	10.95	14.43	13.70
4	0.118	0.125	0	50	50	52	0.00	14.14	14.14	14.43	13.70
5	0.104	0.125	0	80	75	52	0.00	17.89	17.32	14.43	13.70
6	0.090	0.125	0	110	80	52	0.00	20.98	17.89	14.43	13.70
7	0.076	0.076	50	100	100	76	14.14	20.00	20.00	17.44	16.15
8	0.063	0.063	100	75	110	100	20.00	17.32	20.98	20.00	18.61
9	0.049	0.049	150	150	125	124	24.49	24.49	22.36	22.27	21.06
10	0.035	0.035	200	125	150	148	28.28	22.36	24.49	24.32	23.52
11	0.021	0.021	300	250	200	196	34.64	31.62	28.28	27.99	28.43
12	0.007	0.007	400	200	250	244	40.00	28.28	31.62	31.22	33.34
	1	1	1200	1200	1200	1200	161.56	223.31	223.31	229.85	223.31

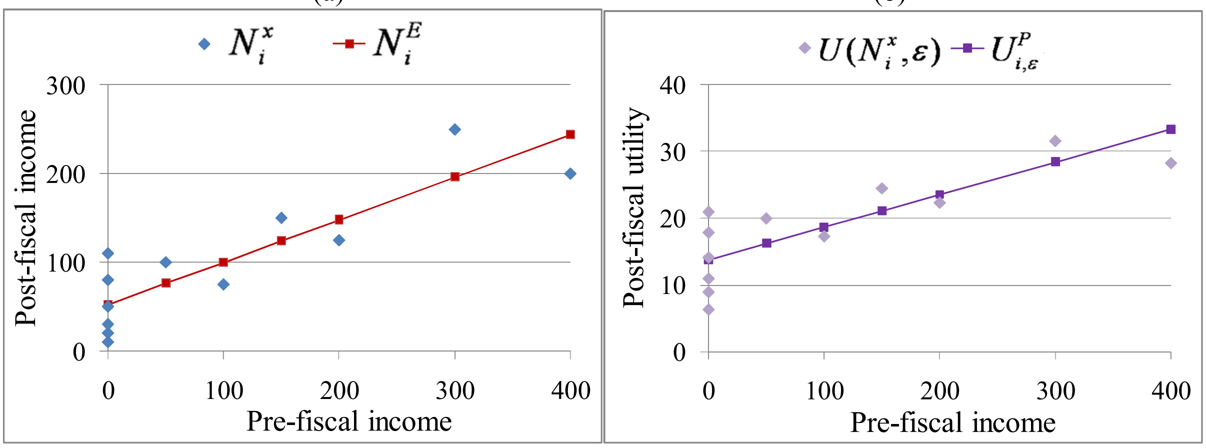
Post-fiscal incomes N_i^x of the first six units are sorted in ascending order, following the procedure from Section 2.1. There is a large variation among incomes within this group of exact pre-fiscal equals. Furthermore, unit #6 has a larger post-fiscal income than units #7 and #8, which is the evidence of reranking; other instances of reranking are between units #5 and #8 (#7 and #8; #9 and #10; #11 and #12).

Table 1 presents different utility vectors obtained for $\varepsilon = 0.5$. Two specific vectors, N_i^E and $U_{i,\varepsilon}^P$, are estimated in the following way: they are calculated as $N_i^E = \widehat{a}^E + \widehat{b}^E X_i^x$ and $U_{i,\varepsilon}^P = \widehat{a}^P + \widehat{b}^P X_i^x$, where $\widehat{a}^E$, $\widehat{b}^E$, $\widehat{a}^P$ and $\widehat{b}^P$ are coefficients obtained by Gini regressions in which X_i^x was an independent variable, while N_i^x and $U(N_i^x, \varepsilon = 0.5)$ were the respective dependent variables.[6]

Figure 1a shows actual post-fiscal incomes N_i^x and expected post-fiscal incomes N_i^E, plotted against the pre-fiscal income X_i^x. Figure 1b presents the utilities of actual post-fiscal incomes $U(N_i^x, \varepsilon = 0.5)$ and their expected values $U_{i,\varepsilon}^P$.

FIGURE 1
EXPECTED POST-FISCAL INCOMES AND UTILITIES

(a) (b)



Notes: Utilities are obtained for $\varepsilon = 0.5$; Gini regressions are estimated for $\nu = 2$.

The inequality indices are calculated for four combinations of parameters ε and ν, and presented in Table 2. The distinction between $\hat{I}(N_i^x)$ and $\widehat{I}(N_i^x)$ enables us to properly capture the reranking effect $\hat{R}^{DJA}(\varepsilon, \nu)$. As (15) reveals, in the presence of pre-fiscal equals, the measure

[6] See Schechtman, Yitzhaki and Artsev (2008) for details about Gini regressions. The beta coefficient for the first regression is $\widehat{b}^E = COV(N_i^x, \widehat{\psi}_i^{x,\nu}) / COV(X_i^x, \widehat{\psi}_i^{x,\nu})$, where $\widehat{\psi}_i^{x,\nu}$ is obtained as $\widehat{\psi}_i^{x,\nu} = s\widehat{\varpi}_i^{x,\nu} / \nu$. The alpha coefficient is $\widehat{a}^E = \hat{\mu}^N - \widehat{b}^E \hat{\mu}^X$.

$\hat{I}(N_i^n)-\hat{I}(N_i^x)$ would underestimate the true amount of reranking by $\hat{R}^{und}(\varepsilon,\nu)$. In our hypothetical case, $\hat{R}^{und}(\varepsilon,\nu)$ is quite high.

When $\varepsilon=0$ and $\nu=2$, we can see that $\hat{C}^{DJA}(\varepsilon,\nu)=0$, which is the result inherent to the DJA model. However, what would happen if the CHI effect was calculated as $\hat{I}(N_i^x)-\hat{I}(N_i^E)$ instead as $\overset{\frown}{I}(N_i^x)-\hat{I}(N_i^E)$? It would be positive and equal to $\hat{R}^{und}(\varepsilon,\nu)$, while the reranking effect would be reduced to $\hat{R}^{res}(\varepsilon,\nu)=\hat{I}(N_i^n)-\hat{I}(N_i^x)$.

TABLE 2
INDICES OBTAINED FOR HYPOTHETICAL POPULATION

	$\varepsilon=0$ $\nu=2$	$\varepsilon=0.5$ $\nu=2$	$\varepsilon=0.5$ $\nu=3$	$\varepsilon=0.5$ $\nu=1$
$\hat{I}(X_i^x)$	0.666667	0.924304	0.984997	0.546834
$\hat{I}(N_i^n)$	0.388889	0.479165	0.619086	0.134222
$\hat{I}(N_i^x)$	0.368056	0.465696	0.607441	0.134222
$\overset{\frown}{I}(N_i^x)$	0.319444	0.411984	0.500583	0.134222
$\hat{I}(U_{i,\varepsilon}^P)$	0.319444	0.411984	0.500583	0.134222
$\hat{I}(N_i^E)$	0.319444	0.349193	0.437733	0.072948
$\hat{\Delta}(\varepsilon,\nu)$	0.277778	0.445139	0.365911	0.412612
$\hat{V}^{DJA}(\varepsilon,\nu)$	0.347222	0.575111	0.547264	0.473886
$\hat{C}^{DJA}(\varepsilon,\nu)$	0.000000	0.062791	0.062850	0.061274
$\hat{R}^{DJA}(\varepsilon,\nu)$	0.069444	0.067181	0.118503	0.000000
$\hat{R}^{und}(\varepsilon,\nu)$	0.048611	0.053711	0.106858	0.000000
$\hat{R}^{res}(\varepsilon,\nu)$	0.020833	0.013469	0.011645	0.000000

3.2. Real data example: Croatian personal taxes and cash social benefits

The fiscal subsystem to be analyzed here consists of three types of social security contributions (SSCs; for the pension, health and unemployment insurance), personal income tax and surtax (PITS), public pensions and six types of cash social benefits.[7]

The data on incomes are taken from the Croatian household budget survey (APK). Since the APK only registers the net incomes of household members, it was a pre-requisite to build a microsimulation model in order to obtain the amounts of PITS and SSCs. The APK is available for years 2001 to 2008, and the samples contain around 3,000 households. The analysis here is based on the 2008 APK sample, consisting of 3,108 households.

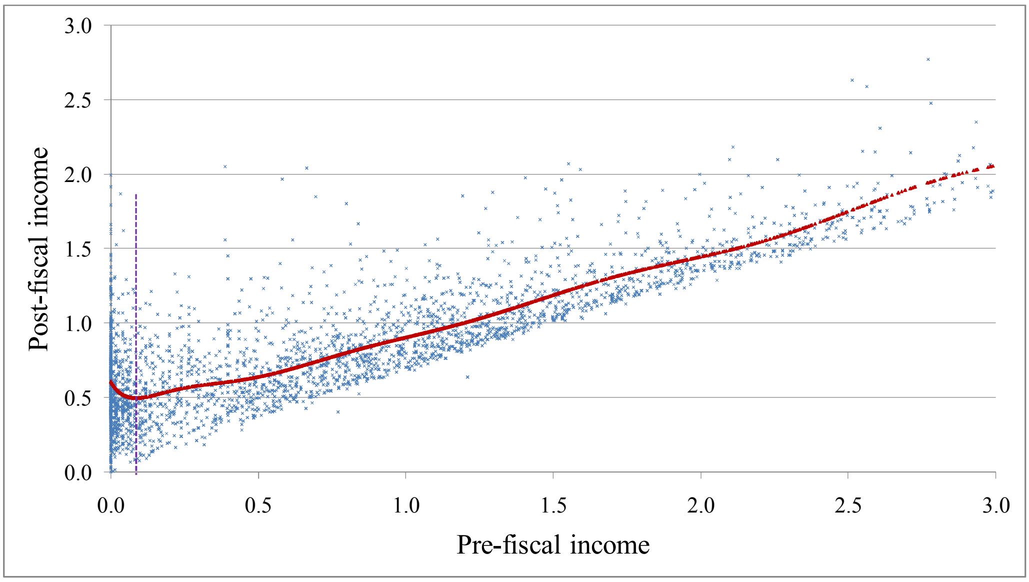
Pre- and post-fiscal incomes are obtained in the following way. Denote with $\dot{X}_k$, $\dot{T}_k$, $\dot{B}_k$ and $\dot{N}_k = \dot{X}_k - \dot{T}_k + \dot{B}_k$, the pre-fiscal income, the sum of all taxes paid, the sum of all benefits received and the post-fiscal income of household k. The incomes are deflated by the equivalence factor obtained using the 'modified OECD scale', $\beta_k = 1 + 0.5(a_k - 1) + 0.3c_k$, where a_k and c_k represent the numbers of adults and children in a household *k*.

Income units are shown in Figure 2, where incomes are expressed in terms of the mean pre-fiscal income. Concentrate on the units to the left of the dotted vertical line; they can be divided into two groups, each covering about 11% of all units, or about 22% in total. In the first group, we have units with a zero pre-fiscal income; in the second group, pre-fiscal incomes are less than 1/10 of the mean pre-fiscal income. On the other hand, the average post-fiscal incomes

[7] Basic support allowances, unemployment benefit, child allowance, sick-leave benefit, maternity and layette supplement, and supplement for the injured and support for rehabilitation and employment of people with disabilities.

for these two groups are 62% and 52% of mean pre-fiscal income, respectively. Thus, for the lower fifth of population, the overwhelming part of income comes from social transfers, mainly public pensions.

FIGURE 2
SCATTERGRAM OF PRE- AND POST-FISCAL INCOMES



EPI vector N_i^E is obtained using the *Curve Fitting Toolbox 1.2* (henceforth: CFT), an interactive tool for graphical data exploration that works within *Matlab R2007b*. CFT enables the use of a dozen of pre-programmed parametric and non-parametric models and provides an opportunity to program one's own model. The model employed in this research is the Fourier series – a sum of sine and cosine functions describing a periodic signal. The number of terms or harmonics chosen is 8 (a maximum allowed by the program), the algorithm used is the Levenberg-Marquardt, and robust fitting options are not used. The top twelve pre-fiscal income units are excluded from the fitting process and their values N_i^E are set to be equal to N_i^x. The identical procedure was used in estimating $U_{i,\varepsilon}^P$.

The advantage of this model is that it can provide us with a spoon-shaped curve, which describes well the specific feature of the current data, where EPI initially falls. The EPI starting point (i.e. when pre-fiscal income is zero) is approximately equal to the mean post-fiscal income for the group of zero pre-fiscal income exact equals. Two other desirable features of this particular method are: (a) the mean of N_i^E is exactly equal to the mean of actual incomes N_i^x; (b) for $\varepsilon = 0$, $\hat{C}^{DJA}(\varepsilon,\nu)$ is very close to zero: for $\nu = 1.2$ it is equal to 0.02% of $\hat{\Delta}(\varepsilon = 0, \nu = 1.2)$, for $\nu = 3$ it is equal to 0.14% of $\hat{\Delta}(\varepsilon = 0, \nu = 3)$. Thus, the estimate convincingly passes the test suggested by Urban (2010), which demands that $\hat{C}^{DJA}(\varepsilon,\nu) \approx 0$ for $\varepsilon = 0$.

The results of the DJA decomposition for the Croatian fiscal system are shown in Table 3. The underestimation of the reranking term, $\hat{R}^{und}(\varepsilon,\nu)$, is relatively small in comparison to RE, but when compared to total HI, measured by $\hat{C}^{DJA}(\varepsilon,\nu) + \hat{R}^{DJA}(\varepsilon,\nu)$, it is not negligible, ranging from 3.3% to 5.6% for different combinations of ε and ν. The meaning of $\hat{R}^{und}(\varepsilon,\nu)$ is explained in the following section.

TABLE 3
INDICES OBTAINED FOR THE REAL FISCAL SUBSYSTEM

	$\varepsilon = 0$ $\nu = 2$	$\varepsilon = 0.5$ $\nu = 2$	$\varepsilon = 0.5$ $\nu = 3$	$\varepsilon = 0.5$ $\nu = 1$
$\hat{I}(X_i^x)$	0.514079	0.702777	0.845047	0.284834
$\hat{I}(N_i^n)$	0.299155	0.340488	0.448383	0.074615
$\hat{I}(N_i^x)$	0.254675	0.299256	0.374014	0.074615
$\overset{\frown}{I}(N_i^x)$	0.252521	0.296900	0.367694	0.074615
$\hat{I}(U_{i,\varepsilon}^P)$	0.252653	0.297038	0.368091	0.074616
$\hat{I}(N_i^E)$	0.252653	0.269239	0.336353	0.053050
$\hat{\Delta}(\varepsilon,\nu)$	0.214924	0.362289	0.396663	0.210219

$\hat{V}^{DJA}(\varepsilon,\nu)$	0.261426	0.433538	0.508693	0.231783
$\hat{C}^{DJA}(\varepsilon,\nu)$	-0.000132	0.027661	0.031341	0.021564
$\hat{R}^{DJA}(\varepsilon,\nu)$	0.046634	0.043588	0.080689	0.000000
$\hat{R}^{und}(\varepsilon,\nu)$	0.002153	0.002356	0.006320	0.000000
$\hat{R}^{res}(\varepsilon,\nu)$	0.044480	0.041232	0.074369	0.000000
$\hat{I}(U^{P}_{i,\varepsilon})-\hat{I}(N^{x}_{i})$	-0.000132	-0.000138	-0.000396	-0.000002
$\hat{R}^{und}(\varepsilon,\nu)$ (% $\hat{\Delta}$)	1.00	0.65	1.59	0.00
$\hat{R}^{und}(\varepsilon,\nu)$ (%HI)	4.6	3.3	5.6	0.00

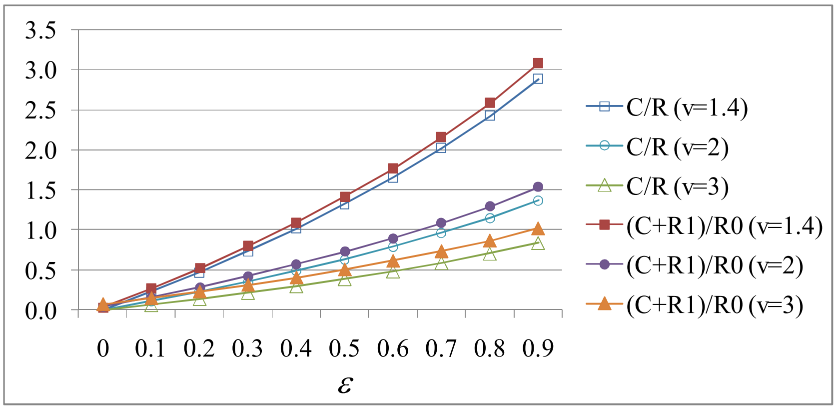
3.3. Reranking effect overestimated

As we have seen, the vertical effect $\hat{V}^{DJA}(\varepsilon,\nu)$ is not affected by the choice of weights. On the other hand, CHI and the reranking effect are, and if the non-adapted weights $\hat{\varpi}_i^{x,\nu}$ are used in the presence of pre-fiscal equals, the former (the latter) effect will be overestimated (underestimated) by the amount of $\hat{R}^{und}(\varepsilon,\nu)$. This irregularity will be obvious in a particular case when $\varepsilon = 0$; here, the DJA CHI effect should be equal to zero by construction. However, if it is calculated as $\hat{C}^{DJA*}(\varepsilon,\nu) = \hat{I}(N_i^x) - \hat{I}(N_i^E)$, it will be positive, which may be confusing to practitioners.[8]

DJA (2003) examined the ratio between the CHI and reranking terms. It should indicate the relative importance of CHI versus reranking in the analysed fiscal system. The ratio of the incorrectly calculated CHI to the reranking indices, $\hat{C}^{DJA*}(\varepsilon,\nu)/\hat{R}^{res}(\varepsilon,\nu)$, will be higher than the ratio of the properly measured ones, $\hat{C}^{DJA}(\varepsilon,\nu)/\hat{R}^{DJA}(\varepsilon,\nu)$. Figure 3 shows these ratios for different values of ν and ε. They increase both with ε and $1/\nu$. The ratio $\hat{C}^{DJA*}(\varepsilon,\nu)/\hat{R}^{res}(\varepsilon,\nu)$ significantly overestimates $\hat{C}^{DJA}(\varepsilon,\nu)/\hat{R}^{DJA}(\varepsilon,\nu)$, by up to 20 percentage points.

[8] If they apply the appropriateness test for the estimate of EPI vector suggested by Urban (2010), they may reject the current EPI estimate although it is the correct one.

FIGURE 3
CHI/RERANKING RATIO

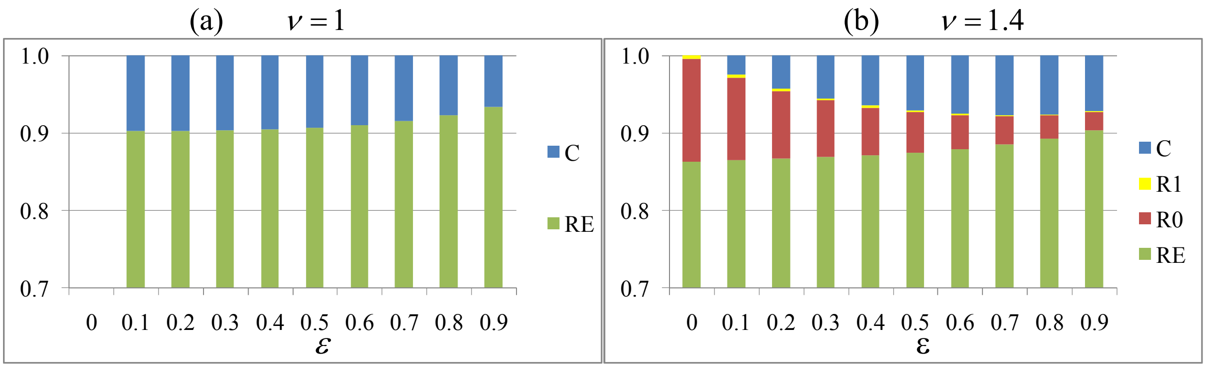


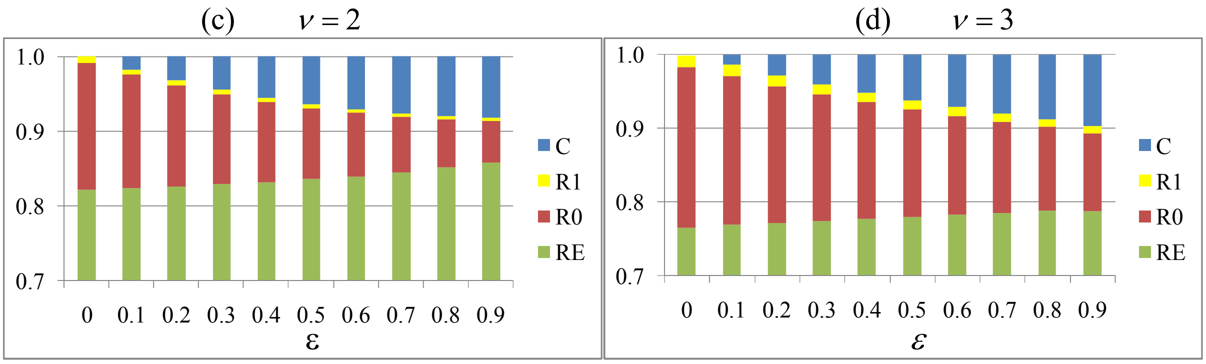
Abbreviations: "C" = $\hat{C}^{DJA}(\varepsilon,\nu)$, "R1" = $\hat{R}^{und}(\varepsilon,\nu)$, "R0" = $\hat{R}^{res}(\varepsilon,\nu)$, "v" = ν

3.4. Structure of the vertical effect

Let us look at the structure of the vertical effect, presented in Figure 4, where all components are separately identified as shares in the total vertical effect. When $\nu = 1$, there can be no reranking; CHI makes about 10% of the vertical effect. If $\nu = 1$ and $\varepsilon = 0$, there is no inequality at all. When $\nu > 1$ and $\varepsilon = 0$, there can be no CHI in the DJA model. For $\nu > 1$ and $\varepsilon > 0$, the ratio between CHI and reranking increases with ε and $1/\nu$, as we have already seen in Figure 3.

FIGURE 4
COMPOSITION OF VERTICAL EFFECT





Abbreviations: "C" = $\hat{C}^{DJA}$, "R1" = $\hat{R}^{und}$, "R0" = $\hat{R}^{res}$, "RE" = $\hat{\Delta}$

3.5. Reranking effect of the counterfactual system

The vertical effect symbolizes the potential RE, the one that would be achieved in the absence of HI. Although the authors conceptualizing the decompositions of RE refrain from suggesting policy makers to eliminate HI, this notion is implicit in the K84, AJL and DJA models. Thus, we can say with certainty that the EPI curve N_i^E of the DJA model eliminates CHI. What about the other manifestation of HI – reranking? Eradication of CHI will nullify reranking, but there are some exceptions. One of them is present in our empirical case. Notice again the shape of EPI curve in Figure 2: it is decreasing on the interval [0,0.1]. It means that some units in this interval have higher expected post-fiscal incomes than some units outside this interval. The counterfactual system (CS) defined by EPI eradicates CHI, but it is not reranking-free. Therefore, a certain amount of HI is implied in the form of reranking.

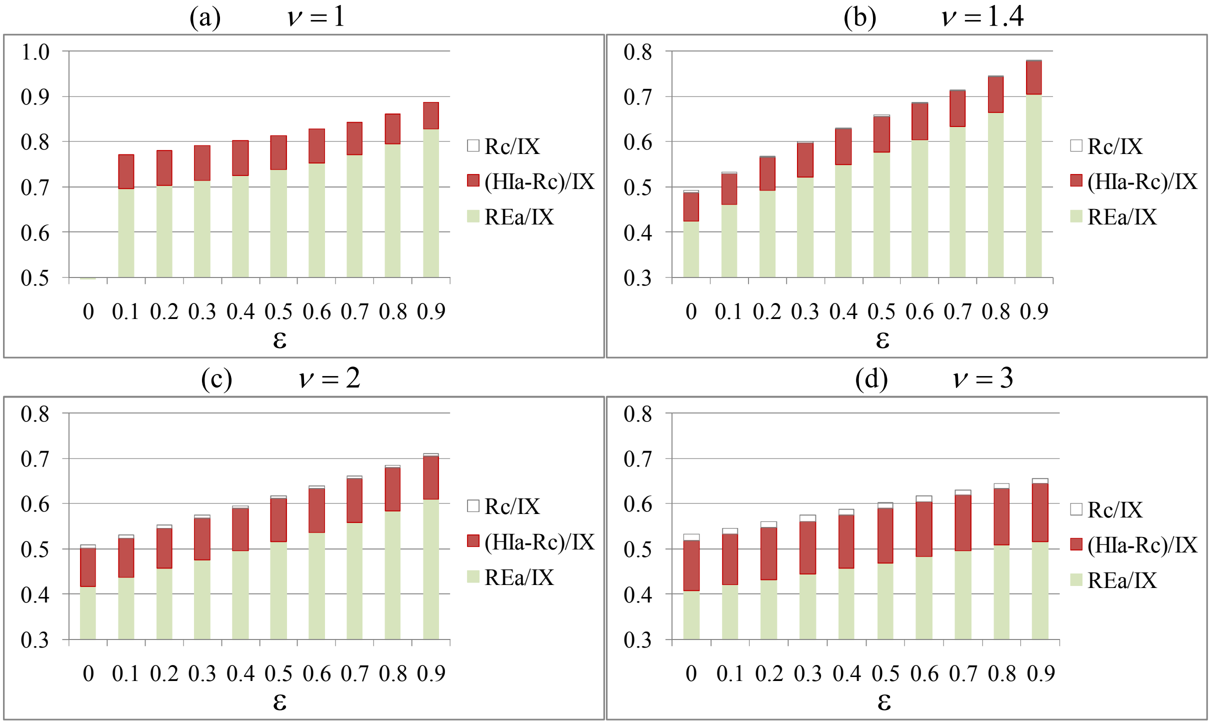
Let us illustrate this on the current example. We can compute the DJA model indicators for the CS as follows. Pre-fiscal incomes and frequency weights are equal to the original ones; $X_i^{x,CS} = X_i^x$ and $\phi_i^{x,CS} = \phi_i^x$. Post-fiscal incomes N_i^x are replaced by expected post-fiscal incomes N_i^E of the actual system; hence $N_i^{x,CS} = N_i^E$. The expected post-fiscal incomes of CS are equal to the expected post-fiscal incomes N_i^E of the actual system; $N_i^{E,CS} = N_i^E = N_i^{x,CS}$. Finally, post-fiscal incomes and frequency weights are sorted together in increasing order to obtain $N_i^{n,CS}$ and $\phi_i^{n,CS}$.

We know in advance that the CHI effect of CS is zero for all ε, because $N_i^{E,CS} = N_i^{x,CS}$. The vertical effect is equal to the one achieved by the actual system because $X_i^{x,CS} = X_i^x$ and $N_i^{E,CS} = N_i^E$. If the vertical effect of the actual system really represents the potential RE, the RE of CS should be equal to $\hat{V}^{DJA}(\varepsilon,\nu)$, the vertical effect of the actual system. However, in an example like ours, this will not be the case. Specifically, we will have that $\hat{I}(N_i^{n,CS}) > \hat{I}(N_i^{E,CS})$; therefore the reranking effect of CS will be greater than zero, $\hat{I}(N_i^{n,CS}) - \hat{I}(N_i^{E,CS}) > 0$, and RE of CS will be $\hat{I}(X_i^{x,CS}) - \hat{I}(N_i^{n,CS}) = \hat{\Delta}^{CS}(\varepsilon,\nu) < \hat{V}^{DJA}(\varepsilon,\nu) = \hat{I}(X_i^{x,CS}) - \hat{I}(N_i^{E,CS})$.

Figure 5 shows different effects that 'constitute' the vertical effect, all of them expressed as shares in the pre-fiscal income inequality. The largest component is RE, $\hat{\Delta}$; next, we have the HI of the actual system reduced by the reranking of CS, $\hat{H}^{DJA} + \hat{R}^{DJA} - \hat{R}^{DJA,CS}$; in the end, there is the reranking of CS itself, $\hat{R}^{DJA,CS}$. The three components together make the vertical effect of the actual system, $\hat{V}^{DJA} = \hat{\Delta} + \hat{H}^{DJA} + \hat{R}^{DJA}$, which supposedly represents the RE that would be achieved if HI were eliminated. However, in our example this is not true: the attainable RE is lowered because EPI induces reranking in the amount measured by $\hat{R}^{DJA,CS}$.

For $\nu = 1$ there can be no reranking and $\hat{R}^{DJA,CS} = 0$. For $\nu > 1$ it increases with ν, from a modest 0.3% of $\hat{I}(X_i^x)$ for $\nu = 1.4$ and $\varepsilon = 0.5$, to 1.4% of $\hat{I}(X_i^x)$ for $\nu = 3$ and $\varepsilon = 0$, when the share of $\hat{R}^{DJA,CS}$ in HI of the actual system reaches 11.4%.

FIGURE 5
COMPOSITION OF VERTICAL EFFECT



Abbreviations: "Rc/IX" = $\hat{R}^{DJA,CS} / \hat{I}(X_i^x)$; "(HIa-Rc)/IX" = $[\hat{H}^{DJA} + \hat{R}^{DJA} - \hat{R}^{DJA,CS}] / \hat{I}(X_i^x)$;
"REa/IX" = $\hat{\Delta} / \hat{I}(X_i^x)$

3.6. Redistributive potential of the fiscal system

Look at the amounts of potential RE; for some value ν, they always increase in ε; when $\nu > 1$, it seems that this increase in ε is faster the lower the parameter ν. More importantly, notice that for some combinations of parameters, potential RE (and actual RE, too) is quite high for the Croatian system of personal taxes, public pensions and cash social benefits. For example, when $\nu = 1.5$ and $\varepsilon = 0.4$,[9] the actual RE reduces the pre-fiscal inequality by no less than 57.6%. The potential reduction is 65.6% if we take into account the reranking effect caused by CF, which is equal to 0.3% of $\hat{I}(X_i^x)$.

[9] The combination of parameters preferred by DJA (2003).

4. Conclusion

The models decomposing redistributive effect into vertical, classical horizontal inequity and reranking effects are highly demanded by practitioners around the world. The Duclos, Jalbert and Araar (2003) decomposition has the prospect to become a leading tool in this area. However, this outlook is possibly jeopardized by difficulties that usually occur during the implementation of the model on the real data. Urban (2010) offers recipes for one of the major problems – the estimation of the expected post-fiscal income curve.

This paper is a follow-up to the said study, dealing with further challenges a researcher may be faced with. More specifically, when there is a large number of pre-fiscal income equals in the sample, the common methods of index computation may produce a misleading estimate of the relative importance of CHI and the reranking effect, and can mistakenly convince the analyst that the estimate of EPI is wrong.

A step-by-step guide for an appropriate application of the DJA model is delivered. The procedure starts with basic data manipulations, continues with the calculation of different income and utility vectors and indices, and ends with the analysis and interpretation of results. The procedure is first employed on a 12-units hypothetical population and then on real data for the Croatian fiscal subsystem comprising social security contributions, personal income tax, public pensions and cash social benefits.

References

Aronson, J. R., Johnson, P. and Lambert, P. J. (1994). ‘Redistributive effect and unequal income tax treatment’, *The Economic Journal*, Vol. 104, pp. 262-270.

Atkinson, A. B. (1970). ‘On the Measurement of Inequality’, *Journal of Economic Theory*, Vol. 2, pp. 244–263.

Atkinson, A. B. (1980). ‘Horizontal Equity and the Distribution of the Tax Burden’, in Aaron, H. J. and Boskins, M. J. (eds), *The Economics of Taxation*, Washington D.C., Brookings, pp. 3-18.

Bilger, M. (2008). ‘Progressivity, horizontal inequality and reranking caused by health system financing: A decomposition analysis for Switzerland’, *Journal of Health Economics*, Vol. 27, pp. 1582–1593.

Duclos, J.-Y., Jalbert, V. and Araar, A. (2003). ‘Classical horizontal inequity and reranking: an integrated approach’, *Research on Economic Inequality*, Vol. 10, pp. 65–100.

Plotnick, R. (1981). ‘A Measure of Horizontal Equity’, *The Review of Economics and Statistics*, Vol. 63, pp. 283-288.

Kakwani, N. C. (1984). ‘On the measurement of tax progressivity and redistributive effect of taxes with applications to horizontal and vertical equity’, *Advances in Econometrics*, Vol. 3, pp. 149-168.

Urban, I. (2010). ‘Decompositions of redistributive effect: practitioners’ support’, *manuscript*.

Yitzhaki, S. (1979). ‘Relative Deprivation and the Gini Coefficient’, *The Quarterly Journal of Economics*, Vol. 93, pp. 321-324