A non-parametric decomposition of redistribution into vertical and horizontal components

Irene Perrote
(Universidad Europea de Madrid)

Juan Gabriel Rodríguez
(Universidad Rey Juan Carlos)

Rafael Salas[1]
(Universidad Complutense de Madrid)

Abstract

We use non-parametric methods to propose horizontal inequity (HI) measures that satisfy certain normative and statistical properties. The HI measures tackle the problem of arbitrary definition of similar individuals and satisfy the horizontal transfer principle. HI is measured by any index consistent with the distance between the estimated and actual post-tax income Lorenz curves. This incorporates an ordinal view of HI. The total effect of a tax system can be decomposed into welfare gain due to income redistribution free of HI and welfare loss due to HI. Other indices in the literature can be seen as particular cases.

JEL Classification: H23, D63, D31 and C14.

Key Words: Vertical redistribution, Horizontal inequity, Social-welfare and non-parametric estimation.

Acknowledgements

This paper has benefited from support from the European Commission Project #ERBCHRXCT980248 and the Spanish Ministry of Education Project #PB98-0546-C0202. We acknowledge useful comments by Valentino Dardanoni, Miguel Ángel Delgado, Peter Lambert and participants in several seminars. Any errors are our own.

[1] Address for correspondence: Rafael Salas, Departamento del Análisis Económico I, Universidad Complutense de Madrid, Campus de Somosaguas, 28223 Madrid (Spain), Phone: +34 91 3942512, Fax: +34 91 3942561, E-mail: R.Salas@ccee.ucm.es

1. Introduction

The principle of horizontal equity states that equals should be treated equally (Feldstein, 1976). Nevertheless, a principle as simple as this has generated a long list of indices to measure the absence of horizontal equity. This is primarily because of the difficulty of finding individuals exactly equal with respect to their equivalent incomes. The principle of horizontal equity should therefore be interpreted in a much less restrictive way to measure the horizontal inequity (HI) from any tax system.

There are two main alternative approaches for measuring HI, each corresponding to a different interpretation of this concept. Reranking analysis focuses on the measurement of the HI induced by the transition from pre-tax to post-tax distribution of income (i.e., Atkinson, 1980, Plotnick, 1981, King, 1983 and Duclos, 1993). Alternative analyses concentrate on the different treatments received by uniform or similar individuals (i.e., Berliant and Strauss, 1985, Aronson et al. 1994, Lambert and Ramos 1997, Camarero et al., 1993 or Pazos et al., 1995). An important deficiency of these indices is the definition of similar individuals, obtained through the division of individuals' incomes into arbitrary intervals. In empirical work, the problem of this ad hoc definition is reduced through a bandwidth sensitivity analysis. This weakness can, however, be solved by non-parametric estimation, as in Lambert and Parker (1997) and Duclos and Lambert (2000).

In this paper we apply non-parametric methods to construct a general index to measure HI with the following characteristics. First, we improve the choice of similar individuals and justify it by the use of a purely statistical criterion that is free from any normative consideration. The identification problem is solved intrinsically by a smoothing technique that optimizes the trade-off between the bias and variance of the estimation (low and large endogenous intervals).

Second, the proposed general index class has normative implications in terms of the so-called *horizontal transfer principle*, and in terms of Lorenz domination that shares similarities with the reranking approach by Atkinson (1980) and Plotnick (1981), and the close-equals approach by Aronson et al. (1994) and Lambert and Ramos (1997). Nevertheless, the reranking approach may be criticized, as reranking is a sufficient, but not a necessary, condition for HI to exist. On the other hand, the close-equals approach uses a particular cardinal representation (Gini coefficient or mean logarithmic deviation) that is unnecessary in our framework. Any (ordinal) S-convex[2] measure can be used in our methodology. Besides, the Aronson et al. (1994) and the "pure" approach by Lambert and Ramos (1997) make the assumption of zero inequity among pre-tax similar individuals. Both assumptions are unnecessary in our setting. Finally, the approach by

Lambert and Ramos (1997) also defines the pseudo-horizontal effect of the tax system, a measure that could be negative, depending on the bandwidth (Lambert and Ramos, 1997, p. 30). In addition, we need not restrict our measurement to homothetic social evaluation functions, as in Duclos and Lambert (2000). Moreover, they do not decompose the overall redistributive effect.

Third, we demonstrate that the Aronson et al. (1994) and the "pure" approach by Lambert and Ramos (1997) are particular cases of our methodology. This is the case when the non-parametric estimation adopts the particular regressogram form over the given exogenous non-overlapping intervals.

A final property of our class of indices that is preserved in our more general context is the additively decomposability of the overall redistributive effect of the tax system into vertical redistribution and HI.

The structure of the paper is as follows. An intuitive graphical idea of the HI is given in Section 2. In Section 3 we suggest how to solve the identification problem of similar individuals by using non-parametric methods. In Section 4, we define the measure of HI, normative properties in terms of Lorenz domination are derived and we decompose the redistributive effect of the tax system into vertical and horizontal

[2] A measure $I: \mathbf{R}^n_{++} \rightarrow \mathbf{R}$ is S-convex if $I(Y) \geq I(AY)$, for all bistochastic matrix A (see

components. Section 5 demonstrates that the Aronson et al. (1994) and Lambert and Ramos (1997) 'pure' approach are particular cases of our methodology. Section 6 consists of conclusions.

2. Graphical notion of horizontal inequity

We start by giving an intuitive idea of the HI notion used throughout this paper.[3] First, we compare Figures 1 and 2 that show the scatter plot of the pre-tax income (X_i) and the post-tax income (Y_i) for the following two polar cases. In Figure 1 we depict the post-tax incomes for a proportional tax system across homogeneous individuals. All individuals are treated equally by this tax system. Zero HI is involved. A clear one-to-one (or functional) relationship between pre- and post-tax income arises in this case. In contrast, in Figure 2 we present the post-tax income values emanating from a tax system with deductions that generates HI. It can be seen how similar individuals receive different fiscal treatments. The one-to-one relationship is no longer satisfied. The key questions are whether or not

Appendix).

there is a one-to-one relationship, and what is the intensity of such a relationship.

We try to answer these questions using a regression curve, as in Figure 3, that presents the estimated (theoretical) value of the post-tax income associated with a given pre-tax income. This is the underlying one-to-one relationship between these variables. The more dispersed are the values of the post-tax incomes, the higher the HI caused by the tax system. Hence deviations from the theoretical curve capture the notion of HI, and the slope of the theoretical curve itself captures the notion of vertical progressivity.

3. Non-parametric solutions to the identification problem

The theoretical function can be approximated either by parametric or non-parametric estimation. While the parametric approach assumes that the estimated curve and, therefore, the tax system have some pre-specified functional form, the non-parametric approach estimates the response without reference to a specific form. This is more realistic in the taxation case where effective marginal tax rates normally do not follow a specific functional form.

The result of the non-parametric smoothing estimation is a continuous function with an almost constant structure in a small neighbourhood of X_i.

---

[3] This idea has been developed in papers, such as Lambert and Parker (1997) and Jenkins and Lambert

The smoothing of the data set consists of calculating a local average of the dependent variable (Y_i) near X_i. Each value of the dependent variable is multiplied by a weight, $\{W(X_i)\}_{i=1...n}$, that may depend on the whole vector $\{(X_i)\}_{i=1...n}$.

Another advantage of the non-parametric estimation is that it overcomes the identification problem. The choice of the size of the intervals for similar individuals depends on the data available and the optimal smoothing technique. This is not an arbitrary criterion, such as dividing the pre-tax equivalent income[4] in a set of groups of individuals (i.e., deciles, centiles or income brackets). The non-parametric estimation optimizes the trade-off between a good approximation to the regression function and a good reduction of the observational noise. If we only minimize the bias of the estimation, interpolating the data produces a large variance. On the other hand, if we minimize the variance through a constant estimation, we have a high bias and large misspecification. The proposed estimator achieves the trade-off between bias and variance.

What does this mean in terms of equity? When the size of the intervals is small, we minimize the bias but increase the probability of considering as different those individuals with rather close incomes. Asymptotically, we

---

(1999).

have that each individual belongs to a different interval and all individuals are different. The HI disappears in this case. However, when the size of the interval is large, we minimize the variance but increase the probability of considering as similar those individuals with rather different incomes. Asymptotically, all individuals are similar and the HI is maximized.

Thus the methodology we propose in this paper minimizes the global error that appears when we create a set of bands that identify similar individuals along the pre-tax equivalent income scale. An arbitrary criterion, such as dividing the pre-tax equivalent income in a set of groups of identical number of individuals, does not minimize the global error. In the next section we show the normative implications of this methodology. We see that a formal definition of HI and some restrictions on the non-parametric technique are required.

4. HI measurement and normative properties

---

[4] This is the common criterion used by the close-equals approach. Alternatively, Jenkins (1988) proposes the use of partitions comprising members sharing the same pre-tax money income and the same type of "fair" characteristics.

In this section, we tackle the central point of the paper, which is to provide an ordinal decomposition of the total redistribution effect of tax RE into vertical redistribution VR and horizontal inequity, such that RE=VR-HI.

Definition: Total Redistribution, Vertical Redistribution and Horizontal Inequity

Given a pre-tax income distribution $X \in \mathbf{R}^n_{++}$, and a post-tax income distribution $Y \in \mathbf{R}^n_{++}$, with the estimated income distribution $Z \in \mathbf{R}^n_{++}$ using the non-parametric technique, we define:

RE(X,Y)=I(X)-I(Y)

VR(X,Y)=I(X)-I(Z)

HI(X,Y)=I(Y)-I(Z)

where I: $\mathbf{R}^n_{++} \rightarrow \mathbf{R}$ is S-convex. HI is measured by an index consistent with the distance between the estimated post-tax income and the actual post-tax income Lorenz curves. Formally, this distance concept can be defined as follows.

Given any pre-tax income distribution X and any post-tax income distribution Y, with an estimated distribution Z, if Z (weakly) Lorenz dominates Y, denoted by $Z \geq_L Y$, then HI(X,Y) ≥ 0. Moreover, given any two post-tax income distributions Y and Y', with different estimated distributions Z and Z', if Y $\geq_L$ Y' and Z' $\geq_L$ Z, then HI(X,Y') $\geq$ HI(X,Y). A fundamental property of this HI definition is that we allow for the use of any (ordinal) S-convex index.[5]

We establish a formal principle that imposes a minimum requirement which we consider any HI definition should satisfy. Our above definition, plus some minimal requirements on the non-parametric estimation, is consistent with this axiom, as seen below. Formally this principle can be stated as follows:

Axiom: the horizontal transfer principle

Consider any pre-tax income distribution X $\in \mathbf{R}^n_{++}$ and any post-tax income distribution Y $\in \mathbf{R}^n_{++}$. Let Y' be another post-tax income distribution, which could be generated from Y by a *horizontal reducing transfer (HRT)* that consists of a fixed amount of income ε>0 between two persons with the same pre-tax income but $Y_i \geq Y_j$, such that $Y_i' = Y_i - \varepsilon \geq Yj' = Yj + \varepsilon$, then HI

5 The vertical and overall redistribution distance concept can be rationalized accordingly. Thus, we can decompose the distance from the Lorenz curve for the pre-tax income (L_X) to the Lorenz curve for the post-tax income (L_Y), which is the total effect of the tax system on the original distribution of income, into two different elements: the vertical redistribution and the HI. The first is the distance from the Lorenz curve for the pre-tax income (L_X) to the Lorenz curve for the theoretical (estimated) post-tax income (L_Z).

should be reduced. On the other hand, a horizontal enhancing transfer, that is $Y_i' = Y_i + \varepsilon \geq Y_j' = Y_j - \varepsilon$, increases HI.

Not all non-parametric estimation methods make the HI measure consistent with this axiom. Our proposal uses the reformulated bistochastic non-parametric estimators proposed in Rodríguez and Salas (2001), where the weights matrix is bistochastic (see Appendix). Any HRT is consistent with the Lorenz dominance criterion. Under any HRT, Y' (weakly) Lorenz dominates Y, and Z'=Z.[6] Then HI measures, based on bistochastic non-parametric estimations, are consistent with the horizontal transfer principle.

There is an obvious link with welfare analysis. Within our framework a horizontal reducing transfer increases welfare, according to any individual social welfare function that is increasing and S-concave. This is so because average incomes of Y, Y' and Z are the same, under the bistochastic non-parametric estimation; and also because of the estimated post-tax income distribution Z does not change, while Y' (weakly) Lorenz dominates the post-tax income distribution Y.

Moreover, due to the bistochastic non-parametric estimation, the estimated post-tax income distribution Z (weakly) Lorenz dominates the

---

6 An innocuous sufficient condition for our proposal to be applied is that the weights are probabilistic (normal and non-negative) on the random variable X. Most of the non-parametric methods in the literature satisfy this condition.

post-tax income distributions Y and Y'. The theoretical (estimated) post-tax income distribution Z, when the tax system does not cause any HI, is always as preferred as the observed post-tax income distribution Y and Y' (see Atkinson (1970) and Dasgupta et al. (1973). Formally, we state the following proposition:

Proposition 1:

Consider any pre-tax income distribution X $\in \mathbf{R}^n_{++}$ and any post-tax income distribution Y $\in \mathbf{R}^n_{++}$. Let Y' be another post-tax income distribution, generated from Y by a *horizontal reducing transfer*, and Z generated by any bistochastic non-parametric estimation, then:

$$W(Z) \geq W(Y') \geq W(Y)$$

for any W(.) individual social welfare function that is increasing and ordinal S-concave.

The proof uses $m_Y = m_{Y'} = m_Z$ and Z $\geq_L$ Y' $\geq_L$ Y. An application of the theorems of Atkinson (1970) and Dasgupta et al. (1973) obtains the result.

Another advantage of the proposed index, implicit in this above proposition, is that HI is always non-negative.[7] Moreover, HI is strictly positive whenever there is any deviation of the actual post-tax income distribution from the estimated post-tax regression curve. Zero HI is guaranteed if both distributions coincide. This was the intuition highlighted in Section 2.

5. Towards a unified framework

In this section we point out the links between the proposed methodology and the alternatives. The normative implications in terms of Lorenz domination are features shared by the reranking approach of Atkinson (1980) and Plotnick (1981) and the close-equals approach of Aronson et al. (1994) and Lambert and Ramos (1997). In subsection 5.1 links with the reranking approach are highlighted, and in subsection 5.2 we explore the connections with the close-equals approach. This is the starting point for a more general and unified framework.

7 An alternative technical explanation of this fact is that Z is the most equitable distribution, in the Lorenz sense, to be obtained using the most favourable set of horizontal reducing transfers. Hence, Z can be seen as the theoretical HI-free distribution.

5.1. Relation between the non-parametric approach and the reranking based HI measures

Reranking-based analysis measures the vertical income redistribution caused by the tax system using the distance between the after-tax income concentration curve ($L_{y,x}$) and the before-tax income Lorenz curve (L_x). Thus, HI is the distance between the after-tax income concentration curve ($L_{y,x}$) and the after-tax income Lorenz curve (L_y).

However, we know from above that the overall redistribution effects can be decomposed under our approach into a vertical redistribution (distance from L_x to L_z) and a horizontal inequity component (distance from L_z to L_y). To this end, both methodologies are analogous (compare Figures 4 and 5). The after-tax concentration curve always lies above the after-tax Lorenz curve in the reranking approach. Similarly, the estimated post-tax income Lorenz curve dominates the post-tax income Lorenz curve in our approach.

Both methods are based on the existence of a theoretical HI-free distribution from which the RE can be decomposed into the VR and HI components. In our case, this benchmark is the non-parametrically estimated L_z, and is the concentration curve $L_{x,y}$ in the reranking approach.

Nevertheless, the reranking approach may be criticized as there may be HI without reranking (see Aronson et al., 1994).

5.2. Close-equals approach as non-parametric modelling

The main decompositions of RE within the close-equals approach can be found in the literature, as in Aronson et al. (1994) and Lambert and Ramos (1997). An interesting question arises: is the use of those approaches an implicit way of applying a non-parametric model? If so, what is the explicit underlying non-parametric model?

To answer these questions, we start with the Lambert and Ramos (1997) 'pure' approach decomposition of the redistributive effect:

$$RE_{LR} = VR - HI \quad (1)$$

$$VR = T_0(X) - T_0^{B,S}(Y) \quad (2)$$

$$HI = T_0^{W,S}(Y) = \sum_{i=1}^{h} \frac{n_i}{N} T_0^{W,i}(Y) \quad (3)$$

$T_0^{B,S}$ and $T_0^{W,S}$ denote the between- and the within-groups Theil-0 inequality indices, evaluated in the usual manner, under the given income brackets disjoint partition S consisting of h exhaustible subgroups. $T_0^{W,i}$ denotes the within-i-subgroup Theil-0 inequality index, where n_i is the population in the i-subgroup and N is the total population. We give an interpretation of this methodology in terms of the theoretical HI-free distribution given in the previous subsection.

According to the definitions in Section 4, the Lambert and Ramos (1997) decomposition is equivalent to a particular econometric model with both vertical redistribution terms equal to:

$$VR = I(X) - I(Z) = T_0(X) - T_0^{B,S}(Y) \tag{4}$$

When we consider the mean logarithmic deviation as the inequality index in our model, the estimated post-tax income inequality and the actual between-groups post-tax inequality are the same, which is guaranteed if the estimated post-tax income equals the mean post-tax income for each close-equals group:

$$Z_i = m_i \ , \quad \forall i \Rightarrow T_0(Z) = T_0^{B,S}(Y) \tag{5}$$

where Z_i is the estimated distribution in the subgroup i and m_i is the mean post-tax income in subgroups. The regressogram (Tukey, 1947) is indeed "an average of the response variables Y of which the corresponding Xs fall into disjoint bins spanning the X observation space" (Härdel, 1990, p. 67). It is, in fact, a kernel estimation (with uniform kernel) evaluated at the midpoints of the bins (Figure 6). Therefore, the Lambert and Ramos (1997) decomposition of RE is equivalent to a particular non-parametric technique: the regressogram.

The same result can be obtained analogously when the horizontal inequity term is considered:

$$
\begin{aligned}
& T_0(Y) - T_0(Z) = \sum \frac{n_i}{N} T_0^{W,i}(Y) \\
& \Leftrightarrow T_0^{B,S}(Y) + \sum \frac{n_i}{N} T_0^{W,i}(Y) - T_0(Z) = \sum \frac{n_i}{N} T_0^{W,i}(Y) \qquad (6) \\
& \Leftrightarrow T_0(Z) = T_0^{B,S}(Y)
\end{aligned}
$$

This relationship can be extended to the decomposition of RE, proposed by Aronson et al. (1994):

$$
RE_{AJL} = G(X) - G^{B,S}(Y) - \sum \frac{n_i}{N} \frac{m_i}{m} G^{W,i}(Y) - R
$$

By using the Gini index (G), Aronson et al. (1994) represents the three terms decomposition of RE above into vertical, horizontal and reranking contributions. $G^{B,S}(Y)$ is the between-groups inequality of the post-tax income, $G^{W,i}(Y)$ is the Gini index of the within-i-subgroup, and R is an index of reranking from the pre- to post-tax income distribution, which is closely related to that of Atkinson (1980) and Plotnick (1981).

Similar conclusions can be obtained if we apply the above procedure to the Aronson et al. (1994) decomposition, using the Gini coefficient.

The decomposition provided by Tukey's non-parametric smoothing achieves the same results as the 'pure' Lambert and Ramos approach (1997) and Aronson et al. (1994) decompositions of RE, depending on the inequality index applied. Therefore, both decompositions of RE are particular cases obtained by the regressogram.[8]

So far we have assumed fixed-width intervals.[9] Nevertheless, windows may well be variable, such as quantiles of the population. *A statistically equivalent block regressogram* can be constructed by averaging over k neighbours. This is, in fact, a k-Nearest Neighbours estimation analogue to the regressogram.[10] The result is a new step function over a different window length.

Two issues emerge when both methods of decomposition are considered as particular cases of the regressogram decomposition of RE. First, the application of a regressogram implies working with a bistochastic matrix of weights.

---

8 The Duclos and Lambert (2000) HI term can also be obtained as an application of Atkinson indices by endogenously generated intervals in the regressogram context.
9 In Lambert and Ramos (1997), bandwidths were 100,000, 50,000 and 10,000 Pesetas per annum.
10 The k-Nearest Neighbours (k-NN) estimate is a weighted average in a varying neighbourhood. When the independent variable is chosen from an equidistant grid, Kernel and k-NN have equivalent weights (Härdle, 1990).

Proposition 2:

A regressogram estimation guarantees that the weights assigned to each observation of the dependent variable sum to one, not only across rows but also across columns; that is, the weights matrix is bistochastic. However, the reverse is obviously not true.

Proof: Let $S=\{s_1(X), \ldots, s_h(X)\}$ be the partition under consideration, $U=\{n_1, \ldots, n_h\}$ the within-groups population set, and $M=\{\mu_1,\ldots,\mu_h\}$ the associate post-tax mean income set. Under the regressogram estimation, we find:

$$z_1^i = \ldots = z_{n_i}^i = m_i, \qquad \forall\, i = 1, \ldots, h \tag{7}$$

In vector notation, Z=BY, where B is the following n-dimensional bistochastic matrix:

$$B = \begin{pmatrix} N_1 & 0 & \Lambda & 0 \\ 0 & N_2 & \Lambda & 0 \\ \Lambda & \Lambda & O & 0 \\ 0 & 0 & \Lambda & N_h \end{pmatrix}$$

and N_i is the n_i-dimensional square matrix:

$$N_i = \begin{pmatrix} 1/n_i & \Lambda & 1/n_i \\ \Lambda & O & \Lambda \\ 1/n_i \Lambda & \Lambda & 1/n_i \end{pmatrix} \quad \forall i = 1,...,h$$

As a consequence of Proposition 2, the estimated post-tax income distribution Lorenz dominates the post-tax income distribution.

Second, the regressogram is one particular non-parametric method. Our method generalizes the current literature. First, we allow for the use of many non-parametric techniques and not only the regressogram, which is considered very simple. Indeed, the regressogram is, by definition, a discontinuous step function that might hide particular features of the distributions within the intervals. To this end, it is worth seeing the difference between the post-tax income estimated according to the regressogram, and the one obtained by the Nadaraya-Watson reformulated estimator, as shown in Figure 7. Even in the regressogram case, we find justification to use not only the Theil 0 index, but any ordinal S-convex index.

6. Conclusions

The central point of this paper is the construction of an ordinal decomposition of total redistribution effect into vertical redistribution and horizontal inequity, which tackles both the normative and statistical issues of the problem. This is done by using non-parametric methods. The indices are based on a generalization of the concept of HI with a different treatment received by adult-equivalent individuals, but it does not have the usual limitation of the existing indices, which is the arbitrary definition of similar individuals. This problem is solved through non-parametric estimation of the post-tax income distribution. The sizes of the intervals depend on the optimal smoothing technique that optimizes the trade-off between the bias (short intervals) and variance (large intervals) of the estimation. Moreover, non-parametric estimation does not take for granted that the tax system adopts any specific functional form.

In addition, the proposed measure has a normative implication that satisfies the horizontal transfer principle. HI is measured by any index consistent with the distance between the estimated and the actual post-tax income Lorenz curves. Because the weight matrix is bistochastic, we prove that the Lorenz curve for the theoretical (estimated) post-tax income (L_Z), when the tax system does not cause any HI, dominates the Lorenz curve for the observed post-tax incomes (L_Y). This allows us to adopt an ordinal versus a cardinal approach, as is common in the literature.

We can also additively decompose the total effect of the tax system on the original distribution of income into two different elements: the first is the welfare gain due to the income redistribution free of HI, and the second is the unambiguous welfare loss due to the HI.

Finally, we demonstrate that the Aronson et al. (1994) and the 'pure' approach of Lambert and Ramos (1997) are particular cases of our methodology. This corresponds to the case when the non-parametric estimation adopts the particular regressogram form over the given exogenous non-overlapping intervals. This suggests an appealing unifying and generalizing framework.

References

Aronson, R., Johnson P. and P. J. Lambert (1994), “Redistributive Effect and Unequal Income Tax Treatment in the U.K.”. *Economic Journal*, 104, 262-270.

Atkinson, A. B. (1970), “On the Measurement of Inequality”. *Journal of Economic Theory*, 2, 244-263.

Atkinson, A.B. (1980), “Horizontal Equity and the distribution of the Tax Burden”, in H. J. Aaron and M. J. Boskin (eds). *The Economics of Taxation*, Washington D.C.: Brookings Institution, 244-263.

Berliant, M. C., and P. Strauss (1983), “Measuring the Distribution of Personal Income Taxes”, in *What role for the Government? Lessons from Policy*

Research. Zeckhauser, R., Leebaert, D. eds., Duke University Press, Durham N.C.

Camarero R., Herrero O. and I. Zubiri (1993), “Medición de la inequidad horizontal: teoría y una aplicación al caso de Vizcaya”. *Investigaciones Económicas*, XVII (2), 333-362.

Dasgupta P., Sen A. and D. Starret (1973), “Notes on the Measurement of Inequality”. *Journal of Economic Theory* 6, 180-187.

Duclos J. Y. (1993), “Progressivity, Redistribution and Equity, with applications to 1985 Britain”. *Public Finance*, 48, 350-365.

Duclos, J. Y. and Lambert P.(2000), “A Normative and Statistical Approach to Measuring Classical Horizontal Inequity”. *The Canadian Journal of Economics.*

Feldstein M. (1976), “ On the theory of Tax Reform”. *Journal of Public Economics*, 6, 77-104.

Härdle W. (1990), *Applied Nonparametric Regression*. Cambridge University Press.

King M. (1983), "An index of inequality: with Applications to Horizontal Equity and Social Mobility”. *Econometrica*, 51, 99-115.

Jenkins, S.P. (1988), “Empirical Measurement of Horizontal Inequity”. *Journal of Public Economics*. 37, 305-329.

Jenkins, S.P. and Lambert P. (1999), “Horizontal Inequity measurement: a basic reassessment” in Silber J. (Eds) *Handbook on Income Inequality Measurement*. Kluwer Academic Publishers.

Lambert P. and Parker S.C. (1997), “Testing for Horizontal Inequity Econometrically”. *Institute for fiscal Studies Working Paper Series 97/18*.

Lambert P. and Ramos X. (1997), “Vertical redistribution and horizontal inequity”. *International Tax and Public Finance*, 4, 25-37.

Pazos M., Rabadán I. and R. Salas (1995), “La desigualdad horizontal en el impuesto sobre la renta de las personas físicas”. *Revista de Economía Aplicada*, 9, 5-20.

Plotnick R. (1981), “A Measure of Horizontal Inequity”. *The Review of Economics and* Statistics, 63, 283-288.

Rodríguez J.G. and Salas R. (2001), “A Bistochastic Nonparametric Estimator”. Mimeo.

Tukey, J. W. (1947), “Non-parametric estimation II. Statistically equivallent blocks and tolerance regions. The continuous case”. *Annals of Mathematical Statistics*, 18, 529-39.

Appendix. Non-parametric estimation

Not all non-parametric estimation methods solve the problem of the normative properties in the measurement of HI, as is mentioned in the text of the paper. Our analysis uses the reformulated bistochastic estimator, proposed by Rodríguez and Salas (2000), which guarantees that the weights matrix is bistochastic. The weights assigned to Y_i sum to one, not only across rows but also across columns.

Given any two-dimensional random sample, $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$, the random variable X denotes the explanatory variable and Y is referred to as the response variable. Our interest is to estimate the regression function at the point *x*,

$$m(x) = E(Y \mid X = x)$$

by a normalized nonparametric estimator; and not only within the intervals, but also between them. We write the nonparametric estimator at point *x* as a weighted average of observations of Y as

$$M(x) = \sum_{j=1}^{n} Y_j W_j(x)$$

The weights W_j can be any probabilistic weights[11] that dampen the Y_js with the corresponding X_j value far from *x*.

However, thus far, the observation weights have only been normalized within the intervals. In our case, we also require the weights to be normalized across the intervals, to achieve the overall convex estimation.

If we estimate the regression function at r distinct points, and the weights matrix is represented by $\mathbf{W}=\{W_{ij}\}_{1\leq i \leq r, 1\leq j \leq n}$, the nonparametric estimator **M** is expressed in vector notation by $\mathbf{M}=\mathbf{W}\cdot \mathbf{Y}$.

An iterative proportional fitting method is applied to the elements W_{ij}, in particular the Deming-Stephan algorithm (Deming and Stephan, 1940), to achieve a normalized matrix of weights across rows and columns, that is a bistochastic one. The algorithm proceeds by row and column adjustments, such that at iteration t (for $\forall$ $t \in \mathbf{N}$), the new elements of the matrix of weights are

$$\overline{W}_{ij}^{(0)} = W_{ij}.$$

If t is odd,

$$\overline{W}_{ij}^{(t)} = \frac{\overline{W}_{ij}^{(t-1)}}{\overline{W}_{i+}^{(t-1)}} = \frac{W_{ij}}{\overline{W}_{i+}^{(t-1)} \times \cdots\cdots \times \overline{W}_{+j}^{(1)} \times \overline{W}_{i+}^{(0)}}.$$

[11] A weight function W_n is said to be a probability weight function if it is normalized ($\Sigma_j W_{nj}(x) = 1$) and nonnegative.

If t is even,

$$\overline{W}_{ij}^{(t)} = \frac{\overline{W}_{ij}^{(t-1)}}{\overline{W}_{+j}^{(t-1)}} = \frac{W_{ij}}{\overline{W}_{+j}^{(t-1)} \times \cdots\cdots \times \overline{W}_{+j}^{(1)} \times \overline{W}_{i+}^{(0)}},$$

where

$$\overline{W}_{+j}^{(t)} = \sum_{i=1}^{r} \overline{W}_{ij}^{(t)}$$

for ∀ j=1, . . ., n; and ∀ t∈**N**

$$\overline{W}_{i+}^{(t)} = \sum_{j=1}^{n} \overline{W}_{ij}^{(t)}$$

and

for ∀ i=1, . . ., r; and ∀ t∈**N**.

Thus the reformulated bistochastic estimator vector is

$$\mathbf{Z} = \mathbf{W}^{\mathbf{B}} \cdot \mathbf{Y},$$

where $\mathbf{W}^{\mathbf{B}}=\{\overline{W}_{ij}\}_{1\leq i \leq r,\ 1\leq j \leq n}$ is the closest bistochastic matrix to **W,** according to the Kullback-Liebler distance function.

Figure 1

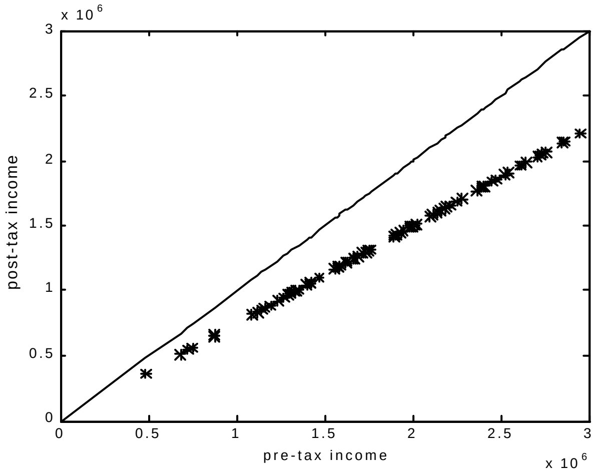

Figure 2

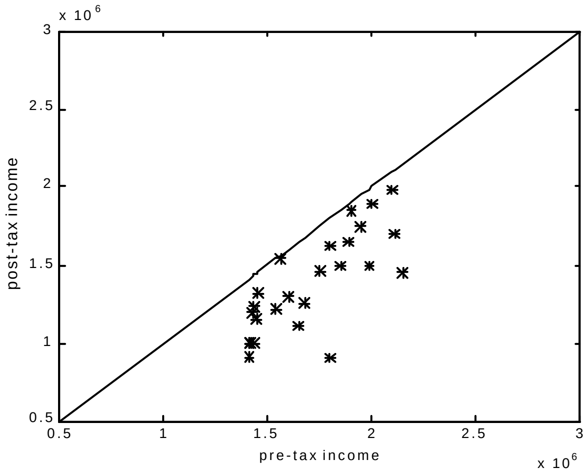

Figure 3

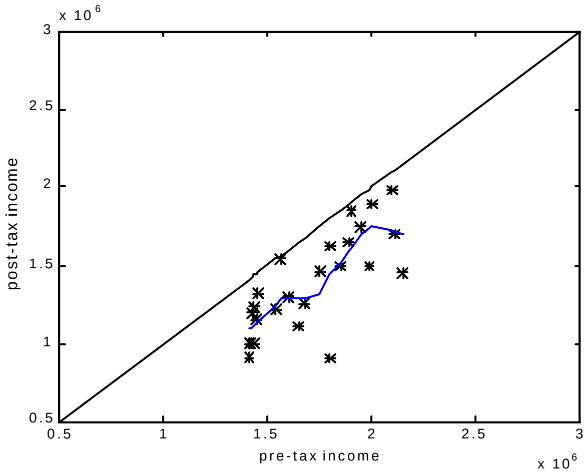

Figure 4

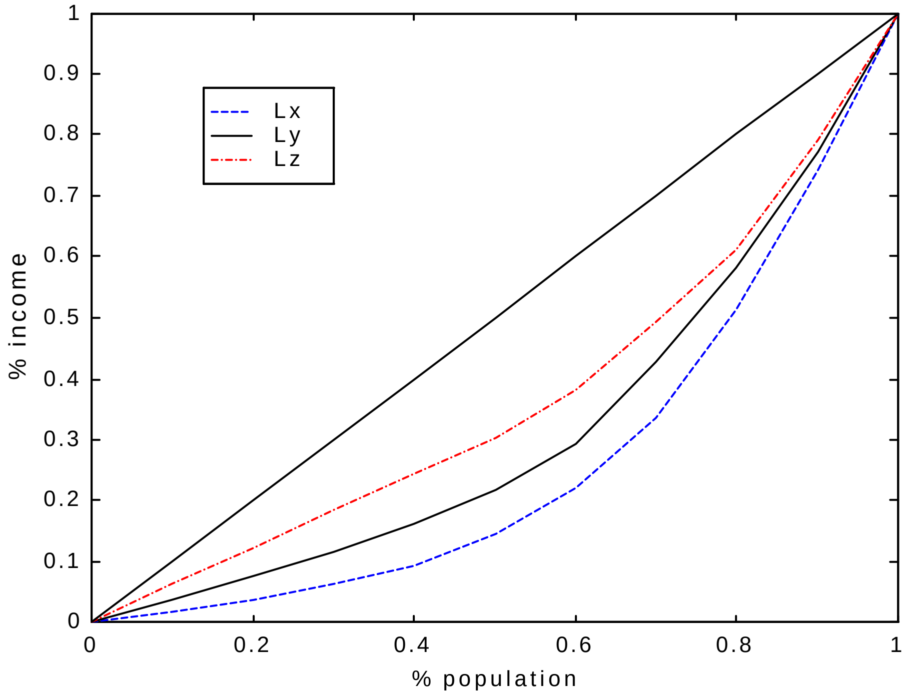

Figure 5

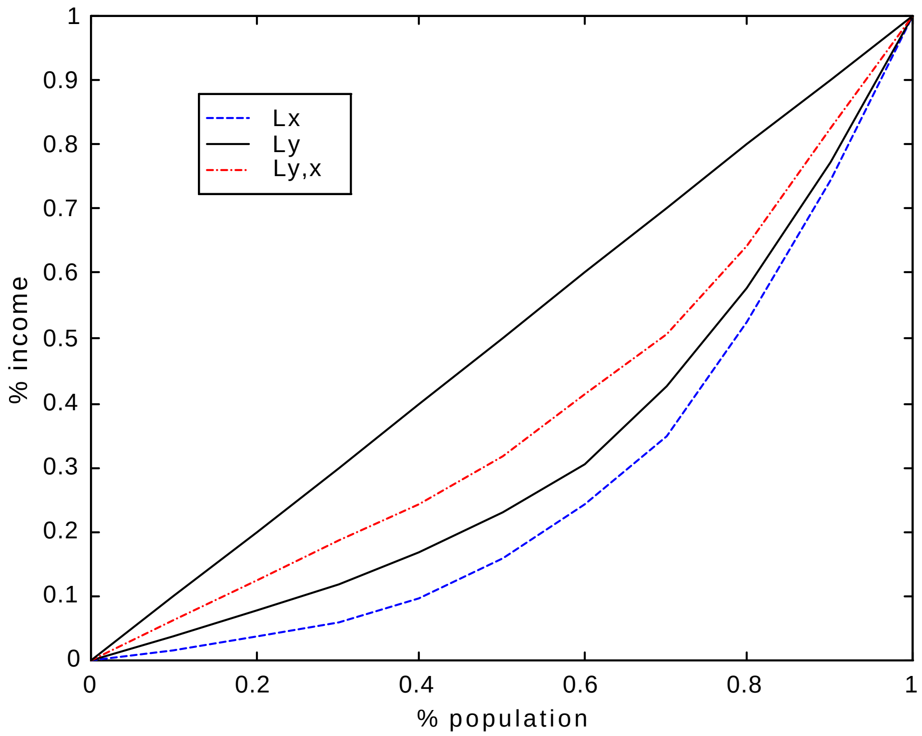

Figure 6

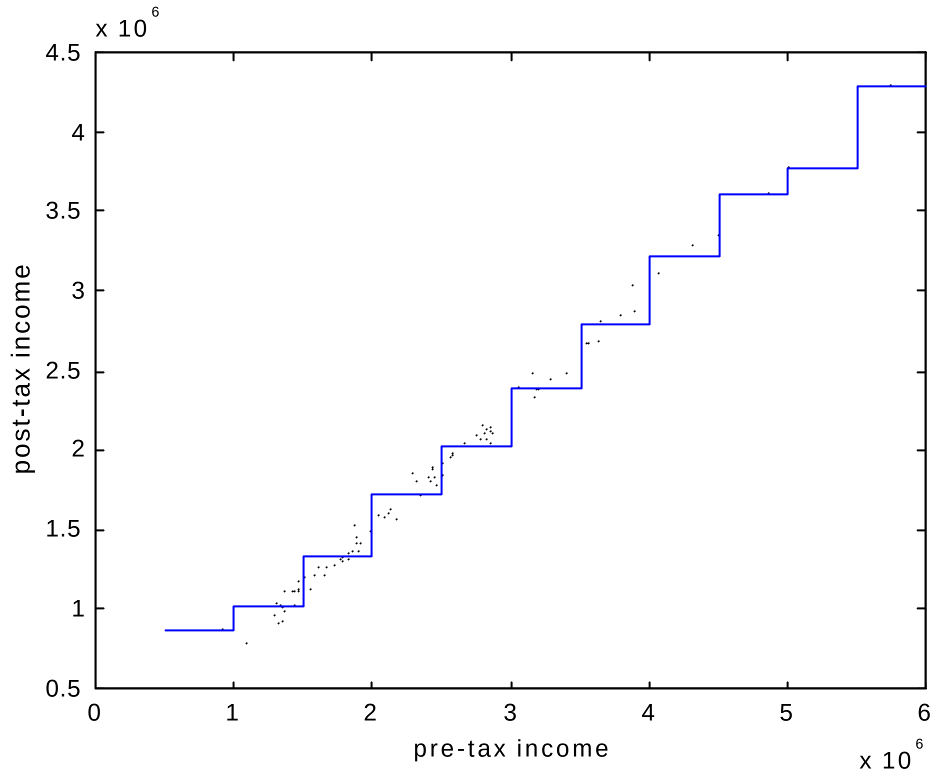

Figure 7