Horizontal Inequity and Vertical Redistribution

PETER J. LAMBERT
University of York, UK

XAVIER RAMOS
ESRC Research Centre on Micro-Social Change, University of Essex, UK and Universitat Autònoma de Barcelona, Spain

Abstract

Inequality of post-tax income among pre-tax equals is evaluated and aggregated to form a global index of horizontal inequity in the income tax. The vertical action of the tax is captured by its inequality effect on average between groups of pre-tax equals. Putting the two together, horizontal inequity measures loss of vertical performance. The identification problem, which has previously been thought insuperable, is addressed by a procedure validating the banding of income units into ‘close equals’ groups. The horizontal and vertical effects of a major Spanish income tax reform are evaluated. Lines for future investigation are suggested.

1. Introduction

Horizontal equity demands that like individuals be treated alike, whilst vertical equity is a command to differentiate appropriately among unlike individuals. These principles can be extended to households and families. We outline a procedure for measuring the extent of horizontal inequity in the personal income tax, leading to a summary index which captures horizontal inequity violations as loss of vertical performance.

The approach we advocate begins by measuring horizontal inequity locally, as inequality of post-tax income among pre-tax equals. Measures of relative and absolute inequality, if used for this purpose, will reflect very different value judgements about the nature and magnitude of an inequity. We select a particular value judgement and local measure, and aggregate into a global index using a particular weighting scheme. This scheme has the property that the importance attributed to a local horizontal inequity does not depend upon the income level at which it is experienced.

The vertical characteristic of the income tax describes the way it treats the unequals. We capture this by the inequality effect of the tax on average between the groups of pre-tax equals located along the income scale. Using the same inequality measure as for local horizontal inequity, an index of vertical redistribution emerges, from which our index of horizontal inequity can be subtracted to obtain the overall redistributive effect of the tax system: horizontal inequity thus gets measured as loss of vertical performance.

The difficulty of identifying the equals in sample microdata makes implementation of any procedure to measure pure (classical) horizontal inequity problematic. Indeed, this difficulty gave rise to the alternative ‘reranking’ approach of Plotnick (1981). However, a variation in our methodology permits the banding of incomes to create ‘close equals’ groups.
A new decomposition of redistributive effect into pseudo-horizontal and pseudo-vertical contributions is thereby obtained. For the interpretation of these pseudo-components, it is as if horizontally, the tax acts to increase inequality within close equals groups and vertically, it acts to reduce inequality between close equals groups. Simulations show that the banding procedure works well. We apply it to Spanish data, to evaluate the horizontal and vertical effects of a major income tax reform.

The organization of the paper is as follows. In Section 2, we define our measures of local and global horizontal inequity and vertical redistribution and provide the central theorem of the paper. In Section 3 we refine the measurement procedure to make it applicable to banded income data or ‘close equals groups’. Section 4 contains the application to the Spanish income tax. In Section 5, we assess the contribution of the paper, showing how, in particular, it responds to a number of points made by other authors in the last 30 years (Johnson and Mayer 1962, Plotnick 1981, Musgrave 1990, Aronson et al. 1994). We also discuss here the nature of the further work that is called for, in order to sustain this new attack on the old problem of measuring horizontal inequity in the income tax.

2. Measuring horizontal inequity and vertical redistribution

When assessing the extent of horizontal inequity in an income tax system, the first step is to turn the business of identifying the equals into a unidimensional problem. We shall require income units’ pre-tax incomes, or living standards, call these \( x \), to be measured on a scale which identifies the equals: \emph{equals will be those having the same pre-tax income} \( x \). Money income would serve for \( x \) if the population under investigation consisted of people with identical tastes and abilities. Econometric modelling of household objectives could yield \( x \) as family indirect utility (Rosen 1978, Manser 1979). More simply, equivalization may be used to provide “a working definition of equity” across family sizes (Gravelle, 1992).

We refer for simplicity to the values \( x \) along this scale as ‘income’. The income unit may be the individual, the family, the household or the ‘equivalent adult’; for convenience we call an income unit a person henceforth. Let \( S(x) \) denote the group of persons having exactly \( x \) before tax: this is the ‘equals group’ located at point \( x \). We want to measure somehow the magnitude of unequal tax treatment among these people. How should we do it? A horizontal inequity is perpetrated whenever two persons in \( S(x) \) experience different losses of ‘income’ from taxation. Let us fix upon units and say that there is a dollar discrepancy between the two. How should this be transformed into a magnitude of inequity?

The approach we take is to measure horizontal inequity in terms of inequality—\emph{inequality introduced by the tax system} where there was none before—that is, within each group of pre-tax equals \( S(x) \). Hence, we focus upon \emph{post-tax inequality among pre-tax equals}. We shall use a measure of relative inequality. This will mean that a 1% deviation in people’s post-tax incomes around the mean in \( S(x) \) gets evaluated as the same amount of horizontal inequity at all pre-tax income levels \( x \). This value judgement is in line with a suggestion of Johnson and Mayer (1962). Were we to use instead a measure of absolute inequality, this would mean that a $1 variation around the mean in \( S(x) \) would count as the
same amount of horizontal inequity wherever it occurred, in line with another of Johnson
and Mayer’s suggestions. We return to this point later.

If and only if there is inequality within $S(x)$, does the tax system display horizontal inequity at income level $x$. The measure we shall use to capture local horizontal inequity is the mean logarithmic deviation (henceforth MLD) of post-tax income, which we denote $J_{S(x)}$ for equals group $S(x)$.\(^2\)

The local measure is now aggregated into a global index. Musgrave (1990) suggested this approach, of devising a local measure and then aggregating: “$HE$ measures which are applicable to particular groups do not suffice. To assess the $HE$ quality of the entire system and to permit comparison with other burden distributions, an overall measure of $HE$ is needed. The construction of such an index is awkward . . . an overall picture [must be] given while inappropriate comparisons between unequals are avoided” (pp. 117–8). We seek to obey Musgrave’s dictum and avoid “inappropriate comparisons” by choosing a weighting scheme to aggregate which is ‘pure’, in that the importance attributed to a local $HE$ violation does not depend upon the income level at which it is experienced.\(^3\)

Our choice of weighting scheme will make the resulting global index commensurate with a measure of the vertical redistribution implicit in the tax system. Let $p_x = N_x/N$ be the population share of the $N_x$ people at point $x$ on the pre-tax income scale (where $N = \sum_x N_x$ is the total population). We aggregate $J_{S(x)}$ using the population shares $p_x$ as weights. The global horizontal inequity index, call it $HI$, is thus defined as:

$$HI = \sum_x p_x J_{S(x)} \quad (1)$$

Let $\mu_{S(x)}$ be the mean post-tax income of the persons in $S(x)$. If the horizontal inequities among these people were to be averaged out, everybody in $S(x)$ would receive $\mu_{S(x)}$. Vertically, the tax operates to compress relative income differentials on average between the groups of pre-tax equals $S(x)$. That is, the vertical action of the tax can be captured through the transformation of each income $x$ in the distribution as follows:

$$x \rightarrow \mu_{S(x)} \quad (2)$$

We define an index of vertical redistribution, $VR$, as the inequality-reducing effect of this transformation:

$$VR = J^b - J^{*b} \quad (3)$$

where $J^b$ denotes inequality in the pre-tax distribution, and $J^{*b}$ denotes inequality in the hypothetical income distribution which would obtain after this ‘averaging out’ of inequities, both measured using the MLD as before. This accords with Musgrave (1990), who also proposes to assess the vertical performance of the income tax in terms of the hypothetical income distribution “assuming the actual distribution among but equal division of the burden within each group of equals” (p. 118).
The central result of the paper is that the two indices HI and VR fully determine the redistributive effect of the tax system:

**Theorem 1.** Let $\text{RE} = J^b - J^a$ be the redistributive effect of the income tax system, where $J^a$ is inequality in after-tax incomes. Then $\text{RE} = \text{VR} - \text{HI}$. 

*Proof:* The MLD is an additively decomposable inequality index. Such an index belongs to the generalized entropy family, and permits a decomposition of overall inequality in any population into a weighted sum of the inequalities present in an exclusive and exhaustive set of subgroups, plus the “inequality between groups”, a contribution which is computed by hypothetically replacing each income in each subgroup by the mean for that subgroup (i.e. eliminating within-group inequality). The weights for the subgroup inequality contributions depend, in general, on both population and income shares. For the MLD, however, the weight is just the population share. Therefore, we may decompose overall after-tax income inequality into its within-equals-groups and between-equals-groups components:

$$J^a = \sum_x p_x J_{S(x)} + J^{aw}$$

(4)

Subtracting this from $J^b$, we have:

$$[J^b - J^a] = [J^b - J^{aw}] - \sum_x p_x J_{S(x)}$$

(5)

which is the result claimed in the theorem, in view of (3) and (1). $\square$

With this result, global horizontal inequity becomes loss of vertical performance. The index HI quantifies the further inequality reduction which could come from eliminating horizontal inequity among every group of equals with no loss in tax yield. The effect of this would be the same as if an income tax schedule

$$T(x) = x - \mu_{S(x)}$$

(6)

had been applied to pre-tax incomes. This notional schedule has the same vertical action as the actual tax because it induces the transformation in (2). Hence the “averaged” schedule $T(x)$ can be thought of as delivering the tax’s vertical performance. The departures from it are the horizontal inequities.5

Because of decomposability of the MLD, the decomposition of RE in Theorem 1 into vertical and horizontal components can be further broken down across socially homogeneous subgroups, and this can be useful if one wishes to avoid using an equivalence scale or to examine the effects of varying the equivalence scale relativities. Suppose, for example, that the population is partitioned into subsets comprising families of given sizes $\kappa = 1, 2, 3, \ldots$. Then we can write overall redistributive effect in the form:

$$\text{RE} = B + \sum \alpha_{\kappa} \cdot \text{VR}_{\kappa} - \sum \alpha_{\kappa} \cdot \text{HI}_{\kappa}$$

(7)
where $\alpha_\kappa$ is the population share in subgroup $\kappa$, $\text{VR}_\kappa$ and $\text{HI}_\kappa$ are the vertical and horizontal measures for subgroup $\kappa$ and $B$ is the between-groups contribution to redistributive effect.\(^b\) The terms $\text{VR}_\kappa$ and $\text{HI}_\kappa$ in (7) are invariant to the scaling of subgroup $\kappa$ incomes, in particular to the equivalence scale used (if any), since they are defined in terms of relative inequality measures. Only $\text{RE}$ and $B$ in (7) are affected by changes in the equivalence scale, the latter revealing the extent to which mean equivalent income differentials between the different family sizes are compressed by application of the tax. In case comparability between families of different sizes using an equivalence scale is eschewed (see Jenkins, 1988, on this), $B$ captures the gap-narrowing effect of the tax system on mean money incomes across the different family sizes.

3. The identification problem

There are typically few exact equals in sample microdata. Pure horizontal inequity will thus be under-estimated if computed from sample microdata. Moreover, to average tax liabilities within equals groups, as for (6), will not lead to a reliable description of the vertical action of the tax system.

Until now, this identification problem has been seen by some authors as an obstacle preventing the measurement of pure horizontal inequity (see e.g. Plotnick, 1981). However, our methodology can be refined to cope with this problem. Specifically, if pre-tax incomes are banded, into ‘close equals groups’, a modified decomposition of redistributive effect can be obtained, into what we may call pseudo-horizontal and pseudo-vertical contributions. For the interpretation of these pseudo-contributions, it is as if horizontally, the tax acts to increase inequality within close equals groups and vertically, it acts to reduce inequality between close equals groups. The details are as follows.

First, select a partition of the pre-tax income distribution into bands. Let $0 = a_1 < a_2 < \ldots < a_{k+1}$ be the values which define the bands (where $a_{k+1}$ is equal or greater than the highest pre-tax income occurring), and let $S_i = \{x : a_i < x \leq a_{i+1}\}$ be the $i$th band. Since we envisage the bandwidths to be small, we may call each $S_i$ a ‘close equals group’ in contrast to the exact equals groups $S(x)$ preceding; formally,

$$S_i = \cup_{a_i < x \leq a_{i+1}} S(x).$$

Extending previous notation, let $J^b_i$ denote inequality of before tax income in band $i$, and let $J^a_i$ denote inequality of after-tax income among the people in band $i$. The increase, at each $i$, is the local ‘pseudo-horizontal’ effect of the tax, and we aggregate this to provide a global index:

$$\text{PHI} = \sum_i p_i [J^a_i - J^b_i]$$

where $p_i$ is the proportion of the population in band $i$. 
For the ‘pseudo-vertical’ action of the tax, let the mean pre- and post-tax income for $S_i$ be $\chi_i$ and $\mu_i$ respectively. If the incomes of the people in each close equals group were to be equalized, both before and after tax, with no change in their total income in either case, each such person would get $\chi_i$ before tax and $\mu_i$ after tax. It would be as if a lump sum tax $\tau(i)$ had been levied on each person according to:

$$\tau(i) = \chi_i - \mu_i$$

(10)

Let inequality before and after this ‘tax’ be $J^{b**}$ and $J^{a**}$ respectively. Our pseudo-vertical index is:

$$PVR = J^{b**} - J^{a**}$$

(11)

This measures the inequality reduction brought about by the actual tax system on average between close equals groups.

**Theorem 2.** Whatever the partition chosen for creating the close equals groups, overall redistributive effect is given by $RE = PVR - PHI$. The finer the partition, the closer $PVR$ and $PHI$ are to $VR$ and $HI$ respectively.

**Proof:** Applying the additive decomposition formula both before and after tax, we obtain

$$J^a = \sum_i p_i J_i^a + J^{a**} \quad & \quad J^b = \sum_i p_i J_i^b + J^{b**}$$

(12)

Now take the difference. As the bandwidths defining the close equals groups get smaller and smaller, $\tau(i)$ approaches $T(x)$ (compare (10) with (6) and use (8)). Hence $PVR$ approaches $VR$, and $PHI$ approaches $HI$.

We can think of ‘pseudo-horizontal inequity’, captured by $PHI$, as a process whereby the incomes of close equals get spread out in the transition from pre- to post-tax; and of ‘pseudo-vertical redistribution’, captured by $PVR$, as a process which narrows relative differentials on average between close equals groups. The ‘spreading out’, picked up by $PHI$ is, of course, net of some ‘compression’, caused by the vertical action of the tax system on relative income differentials within close equals groups. If the bandwidths in the partition $0 = a_1 < a_2 < \ldots < a_{k+1}$ are very wide, this vertical effect could dominate, resulting even in a negative value for $PHI$; one must choose narrow bandwidths to avoid this problem.

To examine the effect of banding and sampling upon the estimates of $VR$ and $HI$, we simulated a distribution of pre-tax income with 10 equals at each of 3,000 income levels, by drawing 3,000 pre-tax incomes $x_i$ from a lognormal distribution (with mean 242 and standard deviation 306) and replicating. Then, we applied a linear income tax schedule with threshold $t = 50$ and marginal rate $m = 25\%$ to each income value $x_i$ (obtaining transformed values $z_i := x_i$ if $x_i \leq 50$, $z_i := 12.5 + 0.75x_i$ if $x_i > 50$), and disturbed the resulting post-tax values $z_i$ for taxpayers only by ten randomly drawn multiplicative factors $v_{ij}(1 \leq j \leq 10)$ taken from a normal distribution with unit mean and standard deviation.
\( \sigma = 0.05 \). This meant that the actual post-tax incomes \( y_{ij} = z_i v_{ij} \) of about 95% of all taxpayers lay within \( 2\sigma \approx 10\% \) of the value \( z_i \) prescribed by the schedule. In this way we obtained 30,000 pre-tax/post-tax income pairs \( (x_i, y_{ij}) \) in terms of which to compute the decompositions of Theorems 1 and 2.

First, we investigated the effect of banding \textit{per se}, using the entire population of 30,000 income units so that no information was lost. (The effect of sampling, to follow, does cause a loss of information and raises statistical issues). How well do the two pseudo-effects \textit{PVR} and \textit{PHI} approximate to \textit{VR} and \textit{HI} using the population data? We found that a uniform bandwidth of 50 (roughly 20\% of mean pre-tax income) was sufficiently wide to cause \textit{PHI} to go negative. In this case there were only 50 close equals groups partitioning the 30,000 income units, plainly too coarse a subdivision. As we narrowed the bandwidth, \textit{PHI} and \textit{PVR} both started to rise, towards their true values, as more and more of the vertical action of the tax took place between (rather than within) close equals groups. See Table 1.

Hence the move from exact to close equals for estimation purposes does not \textit{of itself} cause serious inaccuracy. But this move does not, either, represent a response to the identification problem, which concerns the \textit{loss of information}, in particular about the presence and prevalence of exact equals in the population, which comes \textit{from the act of sampling}. The question is whether, in the presence of sampling, the resort to close equals groups can compensate for the loss of information about exact equals, and lead to estimates for \textit{VR} and \textit{HI} which approximate to the population values. To look into this, we successively selected smaller and smaller random samples from the 30,000-member population. Using a uniform bandwidth of 10 in each case, we computed \textit{PHI} and \textit{PVR} in order to compare these sample estimates with what we know to be the true values: see Table 2.

The results, though merely illustrative, are encouraging: for a bandwidth of 10, and an ever-shrinking sub-population, the estimated decompositions enjoy the same orders-of-magnitude as for the population as a whole. The identification problem, then, should not be viewed as an obstacle preventing the estimation of horizontal inequity from sample microdata. Of course, further investigation is needed here: we have not solved the problem of determining an optimal bandwidth for close equals groups. The bootstrap resampling technique offers the possibility for deeper statistical analysis, which could yield confidence intervals for \textit{PVR} and \textit{PHI} both in terms of sample size and with respect to the bandwidth chosen. Then the sensitivity of results, and in particular the convergence established in Theorem 2 for population data, could be checked with respect to a sample.10

\begin{table}[h]
\centering
\begin{tabular}{cccc}
\hline
Bandwidth & \# groups \( S_i \) & PVR & PHI \\
\hline
50 & 50 & .0677 & neg \\
25 & 85 & .0689 & .0000 \\
20 & 95 & .0691 & .0001 \\
15 & 119 & .0692 & .0003 \\
10 & 159 & .0693 & .0004 \\
5 & 293 & .0694 & .0004 \\
\hline
\end{tabular}
\caption{Narrowing the bandwidth.}
\end{table}

\( \text{the true values: .0694 .0004} \)
The Spanish personal income tax (Impuesto sobre la Renta de las Personas Físicas or IRPF), was until 1988 a tax on family income. A family comprised either a single adult, or a married couple or other extended gathering of adults, with or without children. In 1988, the IRPF was reformed, according to the so-called Ley 20/89 or law 20/89, to permit income splitting within the family; after the enactment of this law, the individual effectively became the tax unit in those families which opted for splitting. Marginal tax rates were also modified in 1988, by the government budget law, and changes in the main tax allowances came into effect as a result of law 6/88.

The ‘extended panel’ of the IRPF is a data set comprising some 200,000 tax declarations by families for each year in the period 1985–1991. Separated tax declarations after the enactment of law 20/89 were combined to produce a single observation for each family. To convert family income into utility, we deflated money income by the OECD’s equivalence scale \( z(A, C) \) for Spain:

\[
z(A, C) = \frac{3 + 7A + 5C}{10}
\]

where \( A \) is the number of adults and \( C \) is the number of children below the age of 14 in the family.

We first examined the trends in inequality and redistributive effect. See Table 3. The inequality trend in pre-tax income is upwards over the period 1985–1990, which was one

<table>
<thead>
<tr>
<th>Sample Size</th>
<th># groups ( S_i )</th>
<th>PYR</th>
<th>PHI</th>
</tr>
</thead>
<tbody>
<tr>
<td>30,000</td>
<td>159</td>
<td>.0693</td>
<td>.0004</td>
</tr>
<tr>
<td>20,000</td>
<td>159</td>
<td>.0693</td>
<td>.0004</td>
</tr>
<tr>
<td>15,000</td>
<td>159</td>
<td>.0694</td>
<td>.0004</td>
</tr>
<tr>
<td>10,000</td>
<td>159</td>
<td>.0690</td>
<td>.0004</td>
</tr>
<tr>
<td>5,000</td>
<td>151</td>
<td>.0681</td>
<td>.0004</td>
</tr>
<tr>
<td>2,000</td>
<td>126</td>
<td>.0703</td>
<td>.0004</td>
</tr>
<tr>
<td>1,000</td>
<td>104</td>
<td>.0676</td>
<td>.0003</td>
</tr>
<tr>
<td>500</td>
<td>92</td>
<td>.0700</td>
<td>.0003</td>
</tr>
</tbody>
</table>

\[ \text{the true values:} \quad .0694 \quad .0004 \]

4. The Spanish income tax

The Spanish personal income tax (Impuesto sobre la Renta de las Personas Físicas or IRPF), was until 1988 a tax on family income. A family comprised either a single adult, or a married couple or other extended gathering of adults, with or without children. In 1988, the IRPF was reformed, according to the so-called Ley 20/89 or law 20/89, to permit income splitting within the family; after the enactment of this law, the individual effectively became the tax unit in those families which opted for splitting. Marginal tax rates were also modified in 1988, by the government budget law, and changes in the main tax allowances came into effect as a result of law 6/88.

The ‘extended panel’ of the IRPF is a data set comprising some 200,000 tax declarations by families for each year in the period 1985–1991. Separated tax declarations after the enactment of law 20/89 were combined to produce a single observation for each family. To convert family income into utility, we deflated money income by the OECD’s equivalence scale \( z(A, C) \) for Spain:

\[
z(A, C) = \frac{3 + 7A + 5C}{10}
\]

where \( A \) is the number of adults and \( C \) is the number of children below the age of 14 in the family.

We first examined the trends in inequality and redistributive effect. See Table 3. The inequality trend in pre-tax income is upwards over the period 1985–1990, which was one


<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( J^0 )</td>
<td>0.2585</td>
<td>0.2977</td>
<td>0.3443</td>
<td>0.3260</td>
<td>0.3454</td>
<td>0.3854</td>
<td>0.3530</td>
</tr>
<tr>
<td>( J^\star )</td>
<td>0.2139</td>
<td>0.2481</td>
<td>0.2886</td>
<td>0.2674</td>
<td>0.2833</td>
<td>0.3211</td>
<td>0.2890</td>
</tr>
<tr>
<td>RE</td>
<td>0.0446</td>
<td>0.0496</td>
<td>0.0557</td>
<td>0.0584</td>
<td>0.0621</td>
<td>0.0643</td>
<td>0.0640</td>
</tr>
</tbody>
</table>
of considerable economic growth. The redistributive effect of the tax is also increasing throughout this period, but not by enough to outweigh the trend in pre-tax inequality, and as a result post-tax inequality is also rising. We cannot detect in the figures for RE an effect from the introduction of splitting in 1988, nor one from the concomitant changes in marginal tax rates and allowances: these two effects are intertwined.

Lasheras et al. (1993) and Castañer (1990) have found similar trends in a Gini coefficient-based measure of redistributive effect, using the same data. Castañer attributed the increase in 1988 entirely to changes in the marginal tax rates and allowances which accompanied the effects of law 20/89, having carried out a simulation to evaluate the impact of the income splitting privilege per se. By comparing their measure for 1988 (with splitting) with the value obtained by simulating a continuance of enforced joint declaration, Castañer concluded that income splitting had hardly had any effect on the redistributive effect of the Spanish income tax system. We may expect the impact of the changes in rates and allowances which accompanied law 20/89 to show in the vertical component VR of our decomposition, and the impact of splitting to come through the horizontal one, HI: the introduction of this privilege was, after all, intended to enhance tax fairness by taking due account of intra-family sharing.

In Table 4 we show the decompositions of RE for 1985–1991 both in level and percentage terms. Thus, in 1985, redistributive effect would have been 101.99% of its actual value if not for the loss caused by horizontal inequity. Vertical redistribution rose in level terms throughout the period, but in percentage terms was declining until 1988, when there was an increase, followed by further reductions. The rise in 1988 can presumably be accounted for by the changes which took place in marginal tax rates and allowances in that year.

It is notable that HI rose to its highest value, both in level and percentage terms, in 1988, following upon the enactment of the splitting provision in law 20/89. It subsequently fell

<table>
<thead>
<tr>
<th>Year</th>
<th>RE</th>
<th>PVR</th>
<th>PHI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985</td>
<td>100.00%</td>
<td>101.99%</td>
<td>1.99%</td>
</tr>
<tr>
<td></td>
<td>0.0446</td>
<td>0.0455</td>
<td>0.0009</td>
</tr>
<tr>
<td>1986</td>
<td>100.00%</td>
<td>101.92%</td>
<td>1.92%</td>
</tr>
<tr>
<td></td>
<td>0.0496</td>
<td>0.0505</td>
<td>0.0009</td>
</tr>
<tr>
<td>1987</td>
<td>100.00%</td>
<td>101.63%</td>
<td>1.63%</td>
</tr>
<tr>
<td></td>
<td>0.0557</td>
<td>0.0566</td>
<td>0.0009</td>
</tr>
<tr>
<td>1988</td>
<td>100.00%</td>
<td>102.49%</td>
<td>2.49%</td>
</tr>
<tr>
<td></td>
<td>0.0584</td>
<td>0.0600</td>
<td>0.0016</td>
</tr>
<tr>
<td>1989</td>
<td>100.00%</td>
<td>102.03%</td>
<td>2.03%</td>
</tr>
<tr>
<td></td>
<td>0.0621</td>
<td>0.0633</td>
<td>0.0012</td>
</tr>
<tr>
<td>1990</td>
<td>100.00%</td>
<td>100.65%</td>
<td>0.65%</td>
</tr>
<tr>
<td></td>
<td>0.0643</td>
<td>0.0648</td>
<td>0.0005</td>
</tr>
<tr>
<td>1991</td>
<td>100.00%</td>
<td>102.30%</td>
<td>2.30%</td>
</tr>
<tr>
<td></td>
<td>0.0640</td>
<td>0.0655</td>
<td>0.0015</td>
</tr>
</tbody>
</table>
again. The impact of law 20/89 was apparently to provoke a one-off exacerbation of horizontal inequity. It may be that law 20/89 promoted changes in pre-tax income distribution which worked to counteract the intended gains in horizontal equity achieved by the tax reform\(^4\)—or it may be that splitting, once fully adopted, introduced a new source of disparity into household income (intra-family income distribution now being relevant). These possibilities are worthy of further investigation: our methodology raises potentially important questions for future tax policy formation.

5. Discussion

The identification problem has been seen as an obstacle preventing the measurement of pure horizontal inequity, because a theoretically sound and operational procedure to identify the equals and measure unequal treatment effects among them was lacking.\(^5\) The reranking approach (Plotnick, 1981) emerged because of this. But the conceptual foundation for replacing the pure classical principle with the no reranking one is unclear (Lambert and Yitzhaki, 1995), and the empirical grounds for it are no longer there since we have provided a way to overcome the identification problem.

Still, reranking remains an interesting phenomenon.\(^6\) Aronson et al. (1994) provide a decomposition of Gini coefficient-based redistributive effect, in which reranking along with a measure of horizontal inequity determine the non-vertical contribution. The link between that paper and ours may be explained as follows. Letting \(I\) be any inequality index, the redistributive effect of the tax system, measured using \(I\), is:

$$\text{RE} = [I^b - I^a]$$

which, because of the identity

$$[I^b - I^a] = [I^b - I^{aw}] - [I^a - I^{aw}]$$

decomposes into two terms. The first term is the redistributive effect of averaged taxes, through the transformation \(x \rightarrow \mu_{SE(i)}\) specified in (5), or vertical redistribution as we have called it. The second term measures the inequality impact of differences in tax treatment. For the MLD, decomposability of \(I^a\) and \(I^{aw}\) across pre-tax equals groups (with pure weights) means that the final term in (15) translates into our index HI. In Aronson et al., (14)–(15) are implemented using the Gini coefficient. In that case, non-decomposability means that the final term in (15) breaks out into a classical horizontal inequity term and a reranking term.\(^7\)

Which decomposition should one use, the Gini-based one or the MLD-based one? One must decide whether rank changes due to taxation matter or social welfare, or not—and opt accordingly. As for the comparative advantages of the respective pure horizontal terms per se, note that Aronson et al.’s, though additive in local Gini coefficients, has income-
dependent (i.e. impure) weights (ibid, p. 265). If the absolute Gini coefficient were used to measure local horizontal inequity and also redistributive effect, the weights would become pure\(^\text{18}\)—and much the same decomposition would emerge. Therefore, the choice between MLD- and Gini-based indices of horizontal inequity can be seen as a choice of inequality concept, retaining purity in weights, or as a choice between pure and impure weights (see on).\(^\text{19}\)

A number of very specific choices guided our development of the indices and decompositions in this paper. First, we chose to measure local horizontal inequity as inequality of post-tax income among equals. Second, we chose to use the relative inequality concept, and third, the mean logarithmic deviation (MLD) measure of this. Fourth, we chose to aggregate local horizontal inequities linearly using pure weights. Fifth, we chose to measure vertical redistribution also using the MLD. Finally, the decomposition of Theorem 1 emerged.

Every step in this sequence should of course be subjected to scrutiny. Musgrave (1990) advocated the local-to-global procedure for measuring horizontal inequity, and warned against “inappropriate comparisons” at the aggregation stage. Our procedures are in line with his suggestions, although he did not argue for using an inequality measure at the local stage. Our choice of the relative inequality concept is in accord with one of Johnson and Mayer’s (1962) suggested value judgements; the absolute inequality concept is in accord with another. Johnson and Mayer’s third proposal, that it might be “worse to inflict one $100 discrimination on one person than to inflict a $10 discrimination on ten people”, might be accommodated within our general framework if other choices of inequality concept and weighting scheme were explored. Our approach to capturing the vertical characteristic of the tax system accords with Musgrave’s line, as well as our horizontal one.

Our analytical framework thus takes forward a line of thought in public finance which had not yet yielded concrete results in measurement theory. Our results show that the approach can be fruitful. An axiomatic development might be the next step. Indeed our results could be seen as the outcome of a set of axioms. Beginning with an axiom to say that overall post-tax inequality should be decomposable in terms of inequality within equals groups and the means and sizes of these groups, which gives local horizontal inequities the same role we have given them and uses the same information for vertical performance as we do (see (2)), one would be led to additive decomposability (Shorrocks, 1984, theorem 1). With the addition of replication and scale invariance, this comes right down to the generalized entropy family (ibid., theorem 5)—of which the MLD is the only member with purity in the weights—and this is our choice.

What kind of value judgement about the nature and magnitude of an inequity could properly be funnelled into the weighting scheme rather than the local measure? An appropriate axiom could direct one to other indices than the MLD for the decomposition of redistributive effect. There is certainly no agreement among the public finance specialists on this issue, and so we do not provide such an axiom. Is it fundamental to capture horizontal inequity as inequality, as we have done? If we are to sustain this new attack on the old problem, it is time the practitioners and the theorists and even the philosophers got together, to decide what is fundamental to horizontal inequity, and to build an axiom system for its measurement de novo.
Acknowledgments

We wish to thank the Editor, Mick Keen, and two anonymous referees for their very helpful comments on earlier versions of this paper. Financial assistance from an EU Human Capital and Mobility Network on ‘The distribution and redistribution of income’ (ER-BCHRXCT940647) is acknowledged.

Notes

1. Later, we shall consider the effect of dropping the assumption of comparability across family sizes using an equivalence scale.
2. In general, the MLD for an income distribution \( \{y_1, y_2, \ldots, y_n\} \) is defined as \( J = \sum f(y_i)/n \). For \( J_{ST} \), the incomes \( y_i \) are the post-tax incomes of those having \( x \) before tax.
3. Global indices of horizontal inequity proposed by Habib (1979) and Berliant and Strauss (1985) involved explicitly income-dependent weighting schemes for individual inequities.
5. This accords with Musgrave (1990): “applied to any one group of equals, HE performance is measured… over what it would have been with equal division of liability within the group” (p. 117).
6. The result in (7) follows by decomposing each of \( J \) and \( J' \) across the subgroups, subtracting to obtain \( \text{RE}_e = B + \sum \alpha_e \text{RE}_{e} \) where \( \text{RE}_{e} \) is the redistributive effect of the income tax in subgroup \( \kappa \), and applying Theorem 1 to each subgroup.
7. Taking the difference in (12) would not work if any other decomposable inequality index than the MLD were used for \( J \): the weights would not be pure population shares in such a case, and would be different in the before and after tax decompositions. It is purity in the weights, a property we have advocated from the outset on normative grounds, which thus permits the identification problems to be solved.
8. If the bandwidth were to be continuously increased, eventually the whole income distribution would be included in a single band. At this point the vertical action of the tax would take place entirely within-band: \( \text{PH} \) would thus approach \( \text{VR} \), whilst \( \text{PV} \) would approach zero (as the between-bands effect vanished).
9. 2,569 of the 3,000 distinct \( x_i \)-values were liable for tax.
10. For an exposition of the bootstrap technique, see Young (1994) and other references cited therein.
11. The growth rates for real GDP at market prices in Spain for the period 1985 through 1990 were: 2.27%, 3.07%, 5.98%, 5.13%, 4.77% and 3.65% respectively.
12. To compute these, we used a uniform set of 1,000 quantiles to define the income bands.
13. See Lasheras et al. (1993) for details and discussion of the changes which took place in allowances and tax rates in 1988.
14. Table 3 certainly shows an appreciable fall in pre-tax inequality in 1988 and subsequent rise. \( \text{RE} \) depends on the distribution of pre-tax income as well as on the tax system. To the best of our knowledge, no research has yet been undertaken to investigate theoretically the consequences for \( \text{RE} \) of changes in pre-tax distribution in the presence of unequal tax treatment effects. This question has been analysed in detail for the case of a progressive income tax schedule (involving no horizontal inequity), using Gini-based measures, by Lambert and Pfahler (1992). Adapting this approach, Lasheras et al. estimate in their paper that for Spain, around 50% of the observed variations in redistributive effect can be “accounted for” by mean income changes, around 11% by tax parameter changes and 37% by changes in the Gini coefficient of pre-tax income.
15. Indeed, Plotnick (1985) complained that any attempt to overcome the identification problem would be “an artificial way to salvage empirical applicability” (p. 241).
17. In Kakwani and Lambert (1995), (14)–(15) are implemented using the Atkinson index. This leads to a money metric welfare measure, having vertical and horizontal components. Kakwani and Lambert employ a specific model of systematic tax discrimination, and do not resort to equals groups or to the measurement of pure/classical horizontal inequity.
HORIZONTAL INEQUITY AND VERTICAL REDISTRIBUTION

18. The horizontal term would take the form \( \sum_r p_r^2 A_r(s) \), where \( A_r(s) \) is the absolute Gini coefficient of post-tax income in equals group \( S(s) \).

19. It is therefore not surprising to find that the Aronson et al. horizontal index gives a very different picture of horizontal inequity in Spain for 1985–1991 than our PHI. The values, of 0.061%, 0.060%, 0.056%, 0.053%, 0.050%, 0.046% and 0.050% do not suggest that law 20/89 had a discernible effect. This observation would guide the policymaker in very different directions from those we have suggested.

References


