6. Measuring the Distribution of Personal Taxes

Marcus C. Berliant and Robert P. Strauss

Abstract

The chapter develops a set of index numbers that compare and aggregate each individual's relative *income and tax position with the income and tax position of other individuals. The index numbers, which are directly viewed as social welfare functions, are then applied to U.S. Treasury Department samples of individual income tax returns to characterize the vertical and horizontal distributions of tax liability. The empirical analysis is performed for 1973, 1975, 1978, and for several widely discussed tax reform proposals. Vertical progressivity is found to increase modestly through time; however, horizontal equity is found to decrease.*

A multiperiod index is also developed which keeps track of each person's relative tax position before and after a tax change, unlike most index numbers. Empirical analysis of this new index indicates significant changes in relative positions of individuals under alternative policies which are not captured by conventional measures. It is suggested that the multiperiod index number be used for policy analysis since before and after relative positions of individuals are of inherent policy interest.

Evaluation of the effects of public policy on distribution or allocation typically requires summarizing large amounts of information. Such aggregation to obtain a "single number" for a busy policymaker usually entails the use of common statistics such as the mean or standard deviation of a critical variable. The normative content of these statistics is usually inferred from their size vis-à-vis that of a current law base case. The prior matter of the *choice* of the particular aggregation rule or statistic is rarely discussed but rather asserted without comment, even though the particular index chosen may be insensitive to wide variations in policy and therefore may not be informative to policymakers.

The use of such aggregation rules or index numbers is quite common in the analysis of taxation and the distribution of income. Although we focus on how to properly characterize the impact of alternative tax policies on the distribution of tax liability, we also raise quite general concerns about how to choose an index number. The indices we have developed can be applied to aggregations of data about a wide range of socioeconomic behavior.

The difference between indices used by policymakers and academic researchers is apparent in the contrast between the measures they use in the area of taxation. The U.S. Treasury Department routinely makes available to the tax committees of the Congress the following indicators: (*a*) the number of taxpayers by income class whose tax liability decreases and the number of taxpayers whose tax liability increases as a result of the policy proposal; (*b*) the average dollar amount of the tax increase or tax decrease by income class; and (*c*) the change in tax burden on representative taxpayers, such as families with various incomes and exemption levels. Such an analysis does not determine whether the tax proposal makes the distribution of after-tax income more equal (regardless of whether more equality is normatively desirable), whether it makes the tax system more progressive vertically, or whether it makes the tax system more horizontally equitable.[1] Another limitation of this type of analysis is that it deals with average effects, although the within-income class variation in tax rates is quite large. For example, the coefficient of variation in effective federal tax rates in 1979 for rather narrowly defined economic income classes ranged from 20 percent to 180 percent.

On the other hand, statisticians and students of income distributions have long used a variety of summary index numbers, such as the Gini coefficient, which are generally more informative than the above qualitative indicators. These index numbers, however, are often computationally burdensome and do not always convey intuitively to policymakers the effects of proposed tax policy changes. For example, a 2 percent change in the after-tax Gini may not be as informative to a policymaker as the statement that 5 percent of all taxpayers experienced tax increases while 35 percent experienced tax decreases, and the remainder experienced no appreciable change.

With regard to the issue of which index number to use, there has been academic interest in deriving index numbers that characterize income distributions from social welfare functions (Atkinson, 1970); Blackorby & Donaldson, 1976), or in finding axioms that are consistent with various index numbers (Kondor, 1975; Fields & Fei, 1978; Sen, 1973; Bourguignon, 1979; and Shorrocks, 1980). Both of these recent approaches attempt to identify more carefully the normative content of various index numbers, although neither approach specifically considers their utilization for public policy analysis.

In choosing among alternative index numbers, one must make several decisions explicitly or implicitly. First, is after-tax income the only variable of interest, or are additional variables significant in determining the equity under scrutiny? If in fact additional variables are of interest, then the index numbers we develop here must be employed because they are naturally *multivariate* in construction in contrast to most conventionally used index numbers, which are *univariate* in construction. On the other hand, if equity is defined in terms of a single variable, one can choose from the usual artillery of index numbers. These univariate index numbers are of several types; however, some are not supported by an axiomatic characterization and are arbitrary in the sense that judgments concerning social welfare are unclear; other univariate index numbers have been at least partially characterized in terms of their axiomatic underpinnings. For example, Kondor (1975) and Fields and Fei (1978) provide conditions that reasonable (univariate) index numbers should satisfy and indicate which index numbers indeed satisfy these axioms. Blackorby and Donaldson (1976), Atkinson (1970), Bourguignon (1979), and Shorrocks (1980) generate certain univariate index numbers axiomatically. The index numbers in this study are also generated by known axioms and are also multivariate in construction.

We seek to develop several new index measures of the horizontal and vertical distributions of income and taxes which are computationally feasible, conceptually complete, and intuitively attractive for policy purposes. To summarize: our principal contributions are: (*a*) the concept that an index number can be viewed directly as an empirical social welfare function, thus obviating the need to deduce it from a social welfare function; (*b*) the creation of a broad class of index numbers, which are based on *relative* comparisons among persons and whose axiomatic underpinnings are completely characterized; (*c*) within this broad class of index numbers, the creation of certain intertemporal index numbers that permit the analysis of equity over time; and (*d*) the extensive empirical implementation of these index numbers for various U.S. tax laws and a variety of widely debated tax reform proposals, using data from the Treasury Department's Individual Income Tax Model. For example, we found an *increase* in the vertical progressivity of U.S. taxes over the period 1973–1979, which is consistent with Bridges (1978) and Okner (1979); however, this was accompanied by significant *decreases* in horizontal equity, which had not been measured previously. Also, the inflation adjustment, usually viewed as proportional in effect, was found to be significantly regressive.

The chapter is organized as follows: section II discusses the choice of an index number; section III develops in detail certain index numbers of vertical and horizontal equity; section IV develops the various tax policy proposals and presents the empirical results; and section V presents conclusions.

On the Choice of Alternative Index Numbers for Evaluation Purposes

Atkinson's Social Welfare Function Approach to Choosing Index Numbers

Sixty years ago, Dalton (1920) pointed out that underlying the choice of one inequality index over another (e.g., choosing the Gini coefficient rather than the variance of income) is some notion of social welfare that would be achieved were an index to reach its limit as a result of incomes' being altered in a particular way. More recently, Atkinson (1970) reiterated this and argued that an index number summarizing the income distribution should be derived from a well-defined social welfare function. Since Atkinson's approach has been the starting point for several recent contributions and differs materially from our approach, we review briefly his line of argument and indicate certain difficulties with it to motivate our justified class of index numbers.

Following Dalton, Atkinson puts certain limitations on the form of social welfare function (SWF) from which he seeks to derive an index number. In particular, he assumes that the SWF is an additively separable and symmetric function of individual income (y). Moreover, Atkinson argues that the reference point in his index number, I, against which the empirical distribution of income should be gauged or compared, is a per capita income or equally distributed income, Y_{EDE}, such that the utility level to society from Y_{EDE} is equal in total to that of the observed distribution of income. The index measure, I, is then stated as:

$$I = 1 - \frac{Y_{EDE}}{\Sigma y / n} = 1 - \frac{Y_{EDE}}{\mu} . \tag{6.1}$$

Based on the work of Pratt (1964) and Arrow (1965), Atkinson indicates that if one assumes that I is invariant to proportional shifts in y, e.g., $I(y) = I(ky)$, $k \neq 0$, then one may deduce that the SWF or $U(y)$ is

$$U(y) = A + B \frac{y^{1-\epsilon}}{1-\epsilon}, \quad \epsilon \neq 1 \tag{6.2}$$

$$U(y) = \log_e(y), \quad \epsilon = 0 ,$$

where ϵ is parametric and represents the degree of aversion to inequality. Atkinson shows that I in discrete terms is

$$I = 1 - \frac{Y_{EDE}}{\mu} = 1 - \left[\sum_i \frac{(y_i^{1-\epsilon})}{\mu} f(y_i) \right]^{\frac{1}{1-\epsilon}} , \tag{6.3}$$

where $f(y)$ is the density of income.

We note five limitations to the approach suggested by Atkinson. First, Atkinson puts certain mathematical limitations on the SWF and, second, certain *other* mathematical limitations on I. That is, he does not solely derive I from a SWF, but deduces I and then imposes certain limitations on I which are imposed on the SWF. From a theoretical point of view, this may be unsatisfactory, for the limitations on I may have implications for the form of the SWF and thereby cancel out a presumed advantage of the approach, namely, that one deduces I solely, and therefore consistently, from a set of initial assumptions or axioms. If I is not solely derived from the SWF, then it is not clear how one decides which axioms are to be placed on the SWF and which are to be placed on I.

For example, the limitation imposed on I which Atkinson needs to obtain equation (6.3) leads to an inconsistent ranking of alternatives as viewed by the SWF, as opposed to I. If the y's increase by $k \cdot y$, $k > 1$, throughout society, then it is apparent from equation (6.1) that social welfare is enhanced. However, I, because it is invariant to proportional shifts, will be indifferent as a measure of well-offness between $I(y)$ and $I(ky)$. Further counter-examples may be developed.

Third, because the axioms on the SWF and I are different, and because they are fundamental value judgments, it is not clear what is being assumed. This blurs the normative underpinnings of the index number and complicates evaluation of it and the normative interpretation of empirical results.

Fourth, the type of derivation Atkinson entertains would appear to limit his derived index numbers to only one variable. This occurs because he derives his index number from the inverse of U, where $SWF = \int U$; the inverse can only be unique for one variable. This is a limitation if one wishes to characterize social welfare in terms of several variables.

A fifth limitation of his derivation of I from SWF is that it is appropriate only for a variable for which U is monotone. For a variable such as income this is, of course, quite reasonable; however, for other variables whose inequality may be of interest,[2] the monotonicity restriction may not be reasonable.

We may also note that while Atkinson requires additive separability of the SWF, I itself is not additively separable. For equation (6.3) to be additively separable in individual utilities, we must be able to separate it so that the social welfare from y_i does not depend on y_j. However, unless $1 - \epsilon$ equals 1 or 0, interaction terms will occur in the expansion of equation (6.3) as the expression is raised to $1 - \epsilon$. Moreover, each y_i is compared in ratio form to μ, which is of course equal to $1/n\ \Sigma_i y_i$ and therefore contains all other persons' incomes. To be sure, Atkinson does not claim that equation (6.3) is or should be additively separable in individual utilities. However, if the resultant I does not have the property of the parent SWF, we question what is gained by placing the restriction on the SWF to begin with.

These considerations raise a question about the proper relationship between social welfare functions and implied index numbers. To the extent that the two are considered separate—the social welfare function reflects abstract concerns while the index number operationalizes these concerns or is more practical in nature—one must place axioms on both, justify each, and relate them to each other.

An alternative approach, which is attractive to the authors, is to consider an empirical index number to be a social welfare function in and of itself. Then only one set of axioms is involved: those that generate the index number. Conversely, one set of axioms may generate an index number directly, rather than relying on one set of axioms for SWF and other axioms on I.

Beyond simplicity and convenience, another reason for viewing an index number as a social welfare function derives from the role such indices can play in public policy. It is common, for example, to measure the degree of progressivity in the tax system through the use of an index number, even though there is no widely accepted concept of progressivity. Higher scores of progressivity, however measured, are often considered more socially desirable. Just as there is no widely accepted view of how to characterize social welfare in its abstract sense, there is frequently no widely accepted view of how to measure empirically such concepts as vertical or horizontal equity. In our opinion, then, social welfare functions and empirical index numbers are similar in that they both refer to subjective notions. Moreover, both social welfare functions and index numbers seek to characterize and rank subjective states of the world. When a concept such as horizontal equity is formalized by a specific mathematical form, subjective judgments are implicitly made which are similar if not identical in form to those made when formalizing a social welfare function.

For these reasons (simplicity and convenience, similarity of normative content, and formal similarity), we treat index numbers as social welfare functions.[3] Below, we describe two general forms of index numbers that meet the five concerns raised with regard to Atkinson's social welfare function approach to index numbers.

A New Class of Group-utility Index Numbers

There are two separate conceptual parts to any index number or social welfare function that one may wish to construct: *(a)* a set of rules that compares values of variable(s) for individuals in the society and as a result creates a "score" or initial index-number value, and *(b)* a set of aggregation rules that combine these individual-level scores to obtain an overall score or level of social utility for the entire society.

Consider, for example, the variance, σ^2, of a distribution of income, y, to be a social welfare function or index number of interest:

$$\sigma^2 = \frac{1}{n}\sum_i (y_i - \bar{y})^2, \text{ where } \bar{y} = \frac{1}{n}\sum_i y_i . \tag{6.4}$$

The comparison rule is $(y_i - \bar{y})^2$, and the aggregation rule is $1/n\ \Sigma$ (i.e., normalized addition). For each ith observation, one creates a comparison or value $(y_i - \bar{y})^2$ and then adds these scores up for all persons in society. Of course, the comparison rule contained in σ^2 is not the only one we might entertain; there are an infinite number of algebraic statements that could be written down.

This characterization of index numbers may be elaborated by examining more closely the nature of the comparison function. In the case of the variance, the comparison is made once for each person's y vis-à-vis $\bar{y}$, and the *form* of the comparison is the squared difference. A more general treatment of the comparison value might be to consider more persons in the comparison, and more variables (e.g., the income and age of each person).

It may appear somewhat unusual to have more than one person in the initial comparison which is then added up; however, Kendall (1947) showed that equation (6.4) can be algebraically transformed so that $(y_i - \bar{y})^2$ becomes

$\frac{1}{2}(|y_i - y_j|)^2$. That is, the accumulation of comparisons contained in the variance between a person's income and the overall mean is equivalent to the accumulation of absolute differences between *pairs* of persons in society. This equivalent statement of the variance must be accumulated across *all possible* pairs of persons in society, whereas the initial statement in equation (6.4) contains an accumulation of individuals vis-à-vis the mean only once. Thus, equation (6.4) tells us that there are n comparisons, whereas the transformed version contains n^2 comparisons.

This equivalent relationship observed between a comparison of all possible pairs of persons' incomes in society and a more traditional statement of an index number suggests that there may be a much richer class of index numbers than is usually considered. Thus, one might argue for expanding the size of the comparison group from two to three, or for choosing a functional form other than squared differences.

In the case of comparisons of pairs of persons in society, a general form of an index number or social welfare function, S, is then

$$S = G\left[C(\underline{y}_1, \underline{y}_2),\ C(\underline{y}_1, \underline{y}_3),\ \ldots\ C(\underline{y}_{n-1}, \underline{y}_n)\right], \tag{6.5}$$

where:

> $\underline{y}$ is the vector of variable(s) of interest for person 1; it could be the income of a person in a n-person society;
>
> C is a comparison function; it could be $\frac{1}{2}|y_1 - y_2|^2$
>
> G is an aggregation function; it could be $\sum_i \sum_j$, $i \neq j$

It is apparent from equation (6.5) that the value of S increases with the value of each group's scores. That is, S is monotone in C. This property of monotonicity may not be desirable in all applications because the value of S is dependent on the initial distribution of y. In other words, S is dependent on the mean of the income distribution, a trait not necessarily desirable in an index number. A more general way of stating this is that the index number has a unit description associated with it (e.g., income). This is not desirable, since units are, in most social science contexts, arbitrary. For this reason and to make the index measure independent of the (mean) level of income, we desire a unitless measure. This leads us to normalize S:

$$S^* = \sigma / \Delta\ ,$$

where σ and Δ are index numbers of form S with the same units.

The treatment of the S-index number as an empirical social welfare function meets all of the concerns raised with respect to Atkinson's derivative approach to obtaining I. First, any initial conditions need only be placed on one object, and the question of consistency in ranking of alternatives between the SWF and I cannot occur. Value judgments that need to be entertained are thus more clearly apparent in the group-utility index approach than in the social welfare function approach. Second, it is obvious that S or S^* can be generalized to more than one variable and variables other than income, whereas I or most variants of it cannot.

Also, in current, related work (Berliant & Strauss, 1980), we have shown that S and S^* may be characterized by a set of consistent initial axioms, and that the axioms are generally sufficient to generate S or S^*.[4] We have already indicated that Atkinson's approach is deficient by contrast, because certain axioms are placed on the SWF, which are not placed on I, and vice versa.

Application of S-index Numbers to the Distribution of Income and Taxes

We now apply the general index numbers developed above to characterize the distribution of taxes for a sample of persons in society. To keep subsequent computations tractable, we compare all possible *pairs* of persons' variable values, and thus fix the group size at 2. Also, because we are interested in the usual subjective notions of vertical and horizontal tax equity, we consider two variables per person in our development below: the pretax, economic income of the ith person, y_i, and the effective tax rate of the person, t_i (the ratio of net taxes to y). We take up first the single time-period case, and then the important two time-period case.

One-period S-index Measure of Taxes and Income

To describe the vertical characteristics of the tax system, we follow Wertz (1978) and partition taxpayers into three groups: The fraction of taxpayers whose liability vis-à-vis others is progressively distributed, ϕ; the fraction of taxpayers whose liability is proportionately distributed vis-à-vis others, θ; and the fraction of taxpayers whose liability is regressively distributed vis-à-vis others, γ, $(\phi + \theta + \gamma = 1)$. A comparison of two taxpayers shows progressivity when both the income and the effective tax rate of one are greater than the income and effective tax rate of the other. Proportionality occurs when the incomes of the two taxpayers are different but the effective tax rates are the same. Regressivity is said to occur when one taxpayer has a larger income but a lower effective tax rate than the other.

To ascertain the extent to which taxes are distributed progressively, proportionately, and regressively, we take into account not only the number for each comparison, but also the degree of the income and tax rate disparity. Our subjective judgment was that it matters whether person A with tax rate of 28 percent and person B with tax rate of 20 percent have similar or very different incomes. Accordingly, we weight each comparison by the absolute difference in income of each pair of taxpayers.

Similarly, it would seem to matter whether the tax rates of A and B are similar or very different. If A had an income of $30,000 and B had an income of $15,000, it would seem important to observe whether their respective tax rates were 28 percent and 20 percent, or 32 percent and 18 percent. The former would appear to be a "less progressive" comparison than the latter. When we account for the disparity in tax rates, we weight by the ratio of tax rates rather than the difference in tax rates for two reasons. First, using the ratio effectively distinguishes between a paired comparison of 14 percent and 10 percent vis-à-vis 54 percent and 50 percent, whereas using (absolute) differences in tax rates would not.[5] Second, using a ratio is more effective mathematically for dealing with proportional comparisons. That is, if $t_i = t_j$ and $y_i \neq y_j$, then $t_i/t_j = 1$ and $|t_i - t_j| = 0$. In the latter case, such weighting would yield an index score of 0 for that comparison, which would be misleading.

Our analysis of tax rates is in terms of effective rates of taxation. Another approach would be to compare individuals in terms of how much income they retain after taxation, or their "after-tax income rate." The two approaches are obviously related. If the effective tax rate is t, then the after-tax income approach to measuring vertical equity involves comparisons of $1 - t$ among taxpayers. The scoring of comparisons in terms of progressivity, regressivity, and proportionality would be the same in both instances, except that progressivity would be deemed to occur when the fraction of retained or after-tax income declined as income rose. Mathematically, $Max\ (t_1/t_2, t_2/t_1)$ and $Max\ (1 - t_1/1 - t_2, 1 - t_2/1 - t_1)$ are monotonically related. Note, however, that the second expression is not invariant to scalar multiplication and thus does not have all the desired properties discussed earlier.

The three fractions (progressive, proportional, and regressive) are obtained essentially by making all possible comparisons among taxpayers, weighting each comparison by the income and tax-rate disparities, and dividing the weighted count of these progressive, proportional, and regressive comparisons by the total number of weighted comparisons.

Horizontal equity, unlike vertical equity, does not admit of progressive, proportional, or regressive distinctions. Usually, horizontal equity denotes identical tax treatment of persons in the same economic circumstances. Measuring horizontal equity thus requires a plausible criterion for testing whether two people's economic circumstances are the same. Whether the absence of the same effective tax rates for persons in the same income class is in a sense "good" or "bad" becomes problematical.[6] Accordingly, we shall measure the extent to which effective rates are *different*, instances of inequity, among all paired comparison of taxpayers, and the extent to which effective tax rates are the same, instances of equity, within each income class. As with the measure of vertical equity, we weight by the ratio of the rank of effective tax rate classes to account for the extent to which horizontal inequity occurs.

A complete, mathematical development of these one-period vertical and horizontal, group-utility index numbers is provided in the appendix (below).

Two-period S-Index Measures of Taxes and Income

The vertical and horizontal index numbers developed above, like other index numbers used for distributional analysis (e.g., the Gini or variance), are static portrayals of the distribution of income and tax burdens among individuals. The group-utility index numbers developed do have the desirable property that each is bounded by 0 and 1, so that one could compare, for example, θ for current law and θ under the proposal. However, both the traditional vertical measure, such as the Gini, or θ, developed above, presume anonymity; that is, the switching of ownership of high and low incomes will not affect the value of the index number when recomputed. For policy purposes this property of anonymity is unsatisfactory, because the policymaker is usually interested in "how different" in distributional impact a tax change will be when compared with current law. These considerations suggest that it would be useful to characterize the relative tax positions of all pairs of taxpayers in society before and after the tax change, and therefore eliminate the anonymity property usually associated with index numbers. Below, we give an intuitive statement of how one may achieve this.

To permit an intertemporal comparison of the relative vertical tax status among pairs of taxpayers, we need to characterize the vertical distribution of taxes in the second period relative to that of the first for each pair of taxpayers in society. "No change" is said to occur if the same *relative* vertical distribution of taxes in the first period is maintained in the second period after the tax change. For example, if initially $y_1 = \$30{,}000$, $y_2 = \$10{,}000$, $t_1 = 0.15$, and $t_2 = 0.05$, we would score that as a progressive comparison in the first period. If t_1 and t_2 remain the same in the second period, then the policy is said to result in "no change" because the *relative* tax rates for the particular individuals did not change. Note that economic income is defined to be independent of tax schemes. We thus characterize as "no change" any maintenance of relative tax position in the second period vis-à-vis the first period, be it progressive, as above, regressive, or proportional.

The characterization of intertemporal progressive and regressive tax changes then follows immediately. If the relative tax position of a pair of taxpayers is more progressive, less regressive, or involves movement from proportionality to progressivity, then the comparison in the second period is said to be more progressive. Similarly, if the relative tax position of a pair of taxpayers in the second time period is less progressive, more regressive, or moves from proportionality to regressivity, then the comparison in the second period is characterized as more regressive.

Table 6.1 displays the various possibilities in period 1 and period 2 and identifies which movements in relative tax position are classed as more progressive, as no change, and as more regressive. Note that every comparison must fit into exactly one category. Once we have decided which comparisons are pro-

Table 6.1. Definition of two-period index number values

Period 1—Initial comparison is		Period 2: More progressive	Period 2: No change	Period 2: More regressive
Progressive	$y_1 > y_2$ $t_1 > t_2$	$\frac{t'_1}{t_1} > \frac{t'_2}{t_2}$	$\frac{t'_1}{t_1} = \frac{t'_2}{t_2}$	$\frac{t'_1}{t_1} < \frac{t'_2}{t_2}$
Proportional	$y_1 \neq y_2$ $t_1 = t_2$	$t'_1 < t'_2$ for $y_2 < y_1$	$\frac{t'_1}{t_1} = \frac{t'_2}{t_2}$	$t'_1 < t'_2$ for $y_1 > y_2$
Regressive	$y_1 < y_2$ $t_1 > t_2$	$\frac{t'_1}{t_1} < \frac{t'_2}{t_2}$	$\frac{t'_1}{t_1} = \frac{t'_2}{t_2}$	$\frac{t'_1}{t_1} > \frac{t'_2}{t_2}$

Note: y is income, persons 1, 2; t is effective tax rate in period 1; and t' is effective tax rate in period 2.

gressive, proportional, and regressive, we can compute the index numbers simply by counting the number of comparisons of each type and dividing the three counts by the total number of comparisons. To see that these index numbers are of form S^*, note that each numerator consists of the sum of paired comparisons: the comparison value is 1 if the comparison is of the proper type and 0 otherwise. Note that the determination of comparison type depends only on three variable values of each member of the pair. Thus, the numerators are of the form S. If every vertical comparison is given a score of 1 irrespective of the variable values, the sum of all comparisons, or the denominator, is of form S. Hence the index number is of the form S^*.

Application of Index Numbers to Alternative Tax Policy Proposals

Data and Policy Proposals

The data bases used for the empirical application of the various index numbers are: 1975 tax data extrapolated to 1978 levels, 1975 tax data, and 1973 tax data; data were supplied by the U.S. Treasury Department.[7] The effective tax rates in our analysis are computed by dividing net taxes due after credits and refunds by our concept of economic income.[8] The empirical results reported here are based on a 112 × 25 matrix of effective tax rates by economic income classes. The tax rates were more finely divided than income to characterize more accurately the vertical aspects of the distribution of income taxes. The 25 economic income classes were chosen so that each class contained roughly 4 percent of all returns.

Five different policy proposals are analyzed in addition to 1978 law at 1978 levels, 1975 law at 1975 levels, and 1973 law at 1973 levels. The proposals analyzed at 1978 levels are: (*a*) 50 percent maximum tax rate on all sources of income; (*b*) taxation of capital gains at ordinary tax rates; (*c*) partial integration

Table 6.2. *S*-index number analysis of 1973, 1975, and 1978 federal tax law (overall)

	1973[a]	1975[b]	1978[c]
Vertical measure			
Progressive	0.882	0.891	0.913
Regressive	0.097	0.095	0.076
Proportional	0.021	0.014	0.011
Horizontal measures			
Equity	0.251	0.166	0.176
Inequity	0.749	0.834	0.824

a. 1973 tax law at 1973 income levels.
b. 1975 tax law at 1975 income levels.
c. 1978 tax law at 1978 income levels.

Source: Computer analysis of Treasury data tapes.

of the corporate and individual income taxes through a flat 133 percent gross-up, 25 percent refundable credit, and repeal of the dividend exclusion; (*d*) a combined package of (*a*)–(*c*); (*e*) a 15 percent inflation adjustment of all nominal tax amounts (exemptions, brackets, etc.).

Empirical Results

Our empirical results are provided first in terms of 1973, 1975, and 1978 tax law, and then in terms of the five policy proposals at 1978 income levels.

Tables 6.2, 6.3, and 6.4 display the one-period vertical and horizontal group-utility index measures for 1973, 1975 and 1978 tax law at their respective income levels. Table 6.2 shows the overall results. In 1973, 88.2 percent of all weighted vertical comparisons among pairs of taxpayers could be characterized as being progressive in character; 9.7 percent of the weighted comparisons were regressive in character, and 2.1 percent of the comparisons were proportional. By 1978, 91.3 percent of all weighted comparisons were progressive (an increase of 3.1 percentage points), and 7.5 percent of the weighted comparisons were regressive. In terms of the vertical characteristics of the distribution of tax liability, 1978 and 1973 are thus rather similar, and the results for 1975 fall between the two.

The horizontal measures, by contrast, reveal substantial differences between 1973 and 1978. In 1973, 25.1 percent of the weighted comparisons were horizontally equitable in character, and 74.9 percent were inequitable. Put another way, within each of the 25 economic income classes, there were three times as many taxpayers with different effective tax rates as there were taxpayers with the same effective tax rates. In 1978 the fraction of taxpayers experiencing horizontal equity dipped to 17.6 percent, or a reduction of 7.5 percentage points—a one-third decline in horizontal equity.

Table 6.3 stratifies the analysis by whether or not the individual taxpayer itemized deductions. Between 1973 and 1978 it is clear that progressivity in-

Table 6.3. *S*-index number analysis of 1973, 1975, and 1978 federal tax law (standard vs. itemized returns)

	1973		1975		1978	
	Itemizers	Standard	Itemizers	Standard	Itemizers	Standard
Vertical						
Progressive	0.837	0.907	0.825	0.901	0.835	0.931
Regressive	0.146	0.064	0.161	0.079	0.153	0.050
Proportional	0.017	0.029	0.014	0.019	0.011	0.019
Horizontal						
Equity	0.082	0.384	0.072	0.217	0.067	0.239
Inequity	0.918	0.616	0.928	0.783	0.933	0.761

creased for those who took the standard deduction and remained about the same for itemizers. Similarly, the decrease in horizontal equity was experienced primarily by those who took the standard deduction. Also of interest is the fact that for any year, horizontal equity was greater for those who took the standard deduction than for those who itemized their deductions.

Table 6.4 stratifies the results by filing status and then by standard or itemized deduction. The general pattern of increased progressivity over time, apparent in table 6.3, is also apparent in table 6.4 for married filing jointly and head of household returns. On the other hand, married filing separately and single returns displayed some decrease in progressivity over time.

Perhaps the most striking result of stratifying by filing status involves the varying degrees of horizontal equity among types of filers. In all three years, single nonitemizers displayed horizontal equity in 64 percent (or more) of the comparisons. In contrast, married filing jointly comparisons were horizontally equitable 19.5 percent of the time in 1973 and 14.5 percent in 1978. The most remarkable increase in horizontal equity occurred for married filing separately returns between 1973 and 1975. In 1973, 29.2 percent of the comparisons displayed horizontal equity, while in 1975, 48.2 percent of the comparisons displayed horizontal equity. This increase in horizontal equity, however, was countered by dramatic decreases for head of household comparisons. Nonitemizers in 1973 displayed equity 23.7 percent of the time; in 1978, only 11.2 percent of the comparisons displayed horizontal equity.[9]

Table 6.5 displays the results for the various proposals at 1978 levels. Several general observations about these results may be made immediately. First, none of the proposals appreciably changes the static vertical measures. The static progressive score hovers around 0.913 to 0.935. It would appear that the progressive structure of the rate schedules, coupled with a high density of taxpayers below the 50 percent marginal rate, ensures that progressivity is maintained. Second, horizontal equity is also reasonably stable among proposals although the refundable, partial integration proposal does substantially decrease horizontal equity.

Table 6.4. *S*-index number analysis of 1973, 1975, and 1978 federal tax law (stratified by filing status and by standard vs. itemized deductions)

	Married filing jointly			Married filing separately			Single			Head of household		
	Total	Itemizers	Non-itemizers	Total	Itemizers	Non-itemizers	Total	Itemizers	Non-itemizers	Total	Itemizers	Non-itemizers
1973 Vertical												
Progressive	0.872	0.855	0.890	0.890	0.805	0.924	0.931	0.772	0.961	0.888	0.830	0.931
Regressive	0.110	0.129	0.086	0.088	0.179	0.047	0.048	0.214	0.014	0.088	0.152	0.042
Proportional	0.018	0.016	0.024	0.022	0.016	0.029	0.022	0.014	0.025	0.024	0.018	0.027
1973 Horizontal												
Equity	0.112	0.093	0.197	0.235	0.118	0.292	0.601	0.106	0.642	0.178	0.109	0.237
Inequity	0.888	0.907	0.803	0.765	0.882	0.708	0.399	0.894	0.358	0.822	0.891	0.763
1975 Vertical												
Progressive	0.908	0.844	0.943	0.861	0.673	0.922	0.890	0.775	0.905	0.944	0.833	0.974
Regressive	0.083	0.143	0.046	0.112	0.313	0.027	0.089	0.213	0.070	0.051	0.155	0.022
Proportional	0.010	0.013	0.010	0.027	0.014	0.051	0.021	0.011	0.025	0.005	0.013	0.004
1975 Horizontal												
Equity	0.085	0.082	0.110	0.372	0.142	0.482	0.599	0.127	0.646	0.084	0.101	0.095
Inequity	0.915	0.918	0.890	0.628	0.858	0.518	0.401	0.873	0.354	0.916	0.899	0.905
1978 Vertical												
Progressive	0.926	0.868	0.953	0.888	0.751	0.937	0.900	0.720	0.957	0.954	0.844	0.977
Regressive	0.067	0.122	0.038	0.093	0.240	0.020	0.080	0.272	0.015	0.042	0.143	0.019
Proportional	0.007	0.010	0.009	0.019	0.009	0.043	0.019	0.008	0.028	0.004	0.013	0.004
1978 Horizontal												
Equity	0.096	0.077	0.145	0.374	0.154	0.486	0.612	0.109	0.665	0.102	0.104	0.112
Inequity	0.904	0.923	0.855	0.626	0.846	0.514	0.338	0.891	0.335	0.898	0.896	0.888

									0.019
						0.028	0.004	0.013	0.004
	0.374	0.154	0.486	0.612	0.109	0.665	0.102	0.104	0.112
0.855	0.626	0.846	0.514	0.338	0.891	0.335	0.898	0.896	0.888

Table 6.5. Vertical and horizontal index numbers at 1978 income levels for alternative tax policies

		Proposal[a]				
Index number	1978 law	50 percent max rate	Capital gains as ordinary income	Partial integration	Combined package	15 percent inflation adjustment
Vertical						
Progressive	0.913	0.911	0.955	0.906	0.931	0.909
Regressive	0.076	0.077	0.033	0.084	0.059	0.078
Proportional	0.011	0.012	0.012	0.010	0.010	0.013
Horizontal						
Equity	0.176	0.176	0.180	0.144	0.145	0.187
Inequity	0.824	0.824	0.820	0.856	0.855	0.613
Vertical two-period						
Progressive	0.000	0.000	0.054	0.081	0.113	0.132
Regressive	0.000	0.021	0.048	0.206	0.207	0.472
No change	1.000	0.979	0.898	0.713	0.680	0.396

a. See text for a full description of each proposal.

If we examine the two-period index numbers, a somewhat stronger pattern of changes may be observed and requires some amendments to the above. For example, as a result of partial integration, 8.1 percent of the comparisons became *more* progressive, and 20.6 percent of the comparisons became *more* regressive than under 1978 current law. Note also that an inflation adjustment, usually viewed as quite neutral, is regressive for better than 47 percent of the weighted comparisons. The reason why the intertemporal results can be so different from the static results is that the static vertical index number is unaltered if a rich taxpayer changes place with a poor taxpayer, while the intertemporal index number will be markedly affected because effective tax rates will change. Such significant changes in the intertemporal, vertical measure under an integration or inflation adjustment regime means that these proposals move significant numbers of taxpayers relative to one another. The results also indicate the utility of keeping track of initial and subsequent positions in the vertical distribution of taxes.

Conclusion

We have found it reasonable to view index numbers as empirical social welfare functions. Such a direct approach to index number construction permits a much more consistent development of the theoretical underpinnings of the index number than the approach suggested by Atkinson (1970). A broad class of index numbers has been developed; they have a variety of desirable theoretical properties.

With these index numbers we have examined the vertical and horizontal characteristics of the distribution of federal individual income tax liability for a number of years. We note here some of our major empirical observations: (*a*) over the period 1973–1975 there was a modest increase in the vertical progressivity of the federal individual income tax. It would appear, however, that this was accompanied by significant deteriorations in horizontal equity; (*b*) examination of a variety of tax proposals, often considered quite radical in their distributional impact, revealed that they did not have significant overall effects on the vertical or horizontal characteristics of the distribution of tax liability; (*c*) a 15 percent inflation adjustment did not affect the vertical and horizontal index measures in any material way; and (*d*) single nonitemizers appear to be subject to the most equitable taxation; (*e*) conclusions (*b*) and (*c*) were overturned (that is, strong vertical effects were found) when the index number accounted for the change in relative tax position of taxpayers under current law and then under the proposal. This measure of before-and-after tax position appears to be quite sensitive to various tax policy changes and may be a useful guide to the effects of complicated tax proposals.

Appendix. Mathematical Development of One-Period Group Utility Index Numbers to Measure Vertical and Horizontal Equity

To facilitate the algebraic development of the index numbers, let there be $i = 1, \ldots m$ ordered, effective tax-rate classes and $j = 1, \ldots n$ ordered, economic income classes for the first group of taxpayers, and let there be $h = 1, \ldots m$ effective tax-rate classes and $k = 1, \ldots, n$ ordered, economic classes of the second group of taxpayers ($i \neq h$, and $j \neq k$, so we do not compare taxpayers to themselves). Further, let N_{ij} be the number of taxpayers in the ijth tax rate–economic income group which is to be compared to N_{hk}, the number of taxpayers in the hkth tax rate–economic income group. Note that increasing subscripts denote higher income and higher effective tax-rate classes, and that $j = k = 1$ is the lowest effective tax-rate class, which empirically will be the lowest-*negative* tax-rate class (among other reasons, because of the refundable earned income tax credit). To deal with a comparison between a positive and negative tax rate, we take a ratio of the tax-rate class *ranks* (or subscripts) rather than the ratio of the average tax rates in the classes themselves.

Of course, any monotone, increasing transformation of tax rates, such as the rank, may be used in lieu of the rates themselves. Thus, negative tax rates may be handled in many ways; how the tax variable enters the index number determines the trade-offs associated with different comparisons. The same reasoning applies to the handling of negative incomes and the manner in which incomes enter into the index number.

We obtain our measure of the extent to which taxes are proportionately distributed, θ, by making all possible comparisons among groups of taxpayers in the same effective tax rate class but with different economic income classes ($j \neq k$), and then add up these proportional comparisons from different effective tax rate classes to get the total number of proportional comparisons. Normalization by the sum of all weighted comparisons, Δ, provides the fraction of weighted comparisons in which tax liability is proportionately distributed:

$$\theta = \frac{1}{\Delta} \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{\substack{k=1 \\ k \neq j}}^{n} \left[N_{ij} \cdot N_{ik} \cdot |Y_{ij} - Y_{ik}| \right] \qquad (6.i)$$

Note that since tax rates are the same in these proportional comparisons, we do not weight by the ratio of rates, since that ratio always equals one.

The fraction of taxpayers whose tax liability is progressively distributed, ϕ, is obtained by accumulating across comparisons in which the effective tax rate and economic income classes of the second group of taxpayers are smaller than those of the first group of taxpayers ($h < i, k < j$), and by accumulating across comparisons in which the effective tax rate and economic income of the second group of taxpayers are greater than the first group of taxpayers ($h > i, k > j$). Since tax rates vary now in these progressive comparisons, we weight by the ratio of the ranks of tax rate classes discussed earlier. Note that in forming the weight for the tax-rate ratio, we always divide the larger rank by the smaller rank of effective tax rates to insure that comparisons are treated symmetrically. Since the first group of progressive comparisons always entails $h < i$, we form the weight as i/h; similarly, since the second group of progressive comparisons always entails $h > i$, we form the weight as h/i:

$$\phi = \frac{1}{\Delta} \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{h<1}^{m} \sum_{k<j}^{n} \left[N_{ij} \cdot \frac{i}{h} \cdot | Y_{ij} - Y_{hk} | \right] +$$

$$\frac{1}{\Delta} \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{h>i}^{m} \sum_{k>j}^{n} \left[N_{ij} \cdot N_{hk} \cdot \frac{h}{i} \cdot | Y_{ij} - Y_{hk} | \right]. \quad (6.\text{ii})$$

The fraction of taxpayers whose tax liability is regressively distributed, γ, is obtained in the same manner as the fraction of taxpayers whose tax liability is progressively distributed, except now $h < i$ and $k > j$ in the first accumulation, and $h > i$ and $k < j$ in the second accumulation. For the comparisons to be regressive, the second group of taxpayers either has lower effective tax rates and greater economic income or higher effective tax rates and lower economic income than the first group of taxpayers. Since in the first accumulation the effective tax rate of the second is lower than the first group of taxpayers, our tax-rate weight for regressivity is formed by i/h. Similarly, our tax-rate weight for the second accumulation is γ. We then have for γ:

$$\gamma = \frac{1}{\Delta} \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{h>i}^{m} \sum_{k<j}^{n} \left\{ [N_{ij} \cdot N_{hk}] \cdot \frac{i}{h} \cdot | Y_{ij} - Y_{hk} | \right\} +$$

$$\frac{1}{\Delta} \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{h>i}^{m} \sum_{k<j}^{n} [N_{ij} \cdot N_{hk}] \cdot \frac{h}{i} \, | Y_{ij} - Y_{hk} | \, . \quad (6.\text{iii})$$

As may be evident, Δ can be obtained from summing the right hand sides of (6.i)–(6.iii) (without the initial $1/\Delta$ terms), or more compactly:

$$\Delta = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{h=1}^{m} \sum_{\substack{k=1 \\ k \neq j}}^{n} \left\{ N_{ij} \cdot N_{hk} \cdot \max \left\{ \frac{i}{h}, \frac{h}{i} \right\} \cdot | Y_{ij} - Y_{hk} | \right\}. \quad (6.\text{iv})$$

If one obtains Δ from equation (6.iv), then γ may be obtained as $1 - \theta - \phi$.

Several comments about the index of vertical tax equity reflected in equations (6.i)–(6.iv) are in order. First, it is invariant to linear transformations of income or tax rates, and is invariant with respect to multiplication or division by a constant of the number of taxpayers. This means that the index is independent of the units of measure. Second, all variations of the numerator and denominator of the index are symmetric and additively separable with respect to comparisons of each of the three types. The index as a whole is invariant with respect to proportional shifts in any factor or factors. Thus, our empirical social welfare function/index number displays variants of the axioms Atkinson recommends for I and SWF.

Recall that the measurement of horizontal equity entails tax-rate comparisons of taxpayers with the same incomes. Thus, since analysis is done within each income class ($j = k$), there are no income differences to weight by. More precisely, we compactly define the fraction of taxpayers with the same income, but whose tax liability is different from other taxpayers with the same income, or the index of horizontal inequity, β, as:

$$\beta = \frac{1}{\delta} \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{\substack{h=1 \\ h \neq i}}^{m} \left\{ N_{ij} \cdot N_{hj} \cdot \max\left(\frac{i}{h}, \frac{h}{i}\right) \right\}, \tag{6.v}$$

where the sum of the inequity and equity comparisons, δ, is:

$$\delta = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{\substack{h=1 \\ h \neq i}}^{m} \left\{ N_{ij} \cdot N_{hj} \cdot \max\left(\frac{i}{h}, \frac{h}{i}\right) \right\} + \sum_{i=1}^{m} \sum_{j=1}^{n} \left\{ N_{ij} \cdot (N_{ij} - 1) \right\}. \tag{6.vi}$$

The second term in equation (6.vi) represents the number of comparisons in which the effective tax rates and income classes are the same ($i = h$), ($j = k$). A total of N_{ij}^2 comparisons are possible; however, this would involve N_{ij} inappropriate comparisons of taxpayers with themselves. Eliminating these cases results in $N_{ij}(N_{ij} - 1)$ comparisons. The complement of β is our measure of horizontal equity. The fractions β and $1 - \beta$ differ from those developed by Wertz (1975), in that the extent of effective tax-rate differences are accounted for in equations (6.v) and (6.vi).