Chapter 4

Relation of Annual to Longer Accounting Periods

In chapter 3 we found that as income decreased from 1929 to 1932, the lower end of the Lorenz curve moved away from the line of equal distribution while the upper end moved toward it. Beginning with 1933 this trend was reversed, and in 1933, 1934, and 1935 the lower end of the Lorenz curve was closer to the line of equal distribution, the upper end further from it than in the preceding year. The Lorenz curves for wages and business profits behaved similarly, becoming less equal with a decline and more equal with an increase in income, while dividends and interest displayed opposite tendencies.

What conclusions can be drawn from these (or similar) changes from one year to the next? Do they constitute random fluctuations around some average annual distribution? Will a series of such annual distributions indicate whether income is becoming more or less equally distributed? Can they be cumulated in such a way that they will indicate what the distribution of income was for the seven years considered as one accounting period? Or, must they be considered only as 'samples', that is, as slices out of a continuing process that reflect the net effects of many forces for a year? And if they are samples, how is the 'universe' they represent to be interpreted?

These and many other puzzling problems are encountered at every turn. In obtaining data on the distribution of income among persons for 12 months, a calendar or fiscal year, what are we trying to measure? Are we content with the knowledge that Fortune smiled on so and so many people? Or do we desire this information as a basis for some judgment concerning the opportunity to make a living existing institutions provide. Do we need both types of information?

A The 'Average' Distribution of 'Annual' Income

Because of its convenience for administrative purposes the calendar year has considerable importance, and annual distributions of income have many uses. Income taxes, for example, are typically based on the calendar year—not because there is any magic about the calendar year but because it seems a good compromise between the government's need for a continuous flow of revenue and the administrative cost of compiling and processing returns. The year is also the customary period used in planning numerous personal and business activities that involve forecasting, such as family budgets, production programs, and sales campaigns.

But the annual distribution of income does not follow a constant pattern. Numerous changes occur from year to year. For some purposes they may safely be ignored, but for others they are of considerable significance. The sample of identical returns affords empirical evidence that they are associated with changes in income but are not evenly distributed throughout the income range. However, in making forecasts based on income distributions, the error arising from not taking account of these year to year changes may be insignificant compared with the margin of error in a forecast of the volume of income. To reduce the labor entailed in forecasting, and in planning activities on the basis of forecasts, an 'average' that described the income distribution during both high and low income years with a fair degree of accuracy would be of value.[1]

[1] One procedure that has been used extensively is to treat the volume of income as an independent variable and correlate some other (dependent) variable with it. It assumes either that for a given volume there is a particular distribution or that differences in the distribution may safely be ignored. See Gardiner Means, *Structure of the American Economy*, Part I (National Resources Planning Board, Washington, D. C., 1939). Cf. *Consumer Expenditures in the United States* (National Resources Committee, Washington, D. C., 1939).

What kind of an average do we want and how can we get it? An average expressed in terms of dollar group intervals, so that we could say that on the average a given number of people received $10,000-15,000 per year would be useful for only the roughest approximations, since relatively small changes in the volume of income might account for a sizeable deviation from the average. A more useful average would be one that eliminated the effects of changes in the volume of income; i.e., an average in terms of 'relative units', e.g., percentages. The Lorenz curve provides us with a convenient method both of displaying graphically the entire distribution in percentage terms and of interpolating to obtain very fine group intervals in dollar units. It is independent of both the volume of income and the number of recipients. An average Lorenz curve, about which the Lorenz curves for a series of annual periods clustered, would thus provide a bench mark for estimating the size distribution of income for any given volume of income and number of recipients.[2]

A Method of Averaging Lorenz Curves

One method of making such an average is to construct a table giving for each percentile, every fifth percentile, or some other convenient interval of the number of persons, the cumulated percentage of income they received in each of the years to be averaged. The data could be obtained from smoothed Lorenz curves. The percentages could be translated into dollars by multiplying them by the total income for the year to which they relate. By addition, the amount of income received by the lowest 5 percent, the lowest 10 percent, etc., could be ascertained for all the years included in the average. Expressing these amounts as percentages of income for the entire period furnishes the data requisite to construct a Lorenz curve, which is the average for the period (Table 11).

Such an average is a weighted average of the *annual* distribu-

[2] One method of translating the Lorenz curve into a distribution in dollar intervals is described by Charles Merwin, American Studies of the Distribution of Wealth and Income by Size, *Studies in Income and Wealth,* Vol. Three (National Bureau of Economic Research, 1939), p. 16.

Table 11

Method of Averaging Lorenz Curves
Net Taxable Income, 1929-1930
Sample of Identical Taxpayers, 1929-1935

% of families cumulated from lowest	% of net taxable income cumulated from lowest		Net taxable income cumulated from lowest ($000)			% distribution of 1929 + 1930 income cumulated from lowest
	1929	*1930*	*1929*	*1930*	*1929 + 1930*	
5	1.0	0.9	307	251	558	1.0
10	2.5	2.3	767	642	1,409	2.4
15	4.4	4.2	1,349	1,173	2,522	4.3
20	6.8	6.6	2,085	1,843	3,928	6.7
25	9.5	9.2	2,913	2,568	5,482	9.4
30	12.5	12.2	3,834	3,406	7,239	12.4
35	15.9	15.5	4,876	4,327	9,203	15.7
40	19.6	19.1	6,011	5,332	11,343	19.4
45	23.2	22.9	7,115	6,393	13,508	23.1
50	27.2	26.8	8,342	7,482	15,823	27.0
55	31.1	30.9	9,538	8,626	18,164	31.0
60	35.4	35.0	10,856	9,771	20,627	35.2
65	39.8	39.4	12,206	10,999	23,205	39.6
70	44.4	44.2	13,617	12,339	25,956	44.3
75	49.3	49.5	15,119	13,819	28,938	49.4
80	54.9	55.2	16,837	15,410	32,247	55.0
85	61.0	61.7	18,707	17,225	35,932	61.3
90	68.0	69.0	20,854	19,263	40,117	68.5
95	76.7	78.2	23,522	21,831	45,353	77.4
100	100.0	100.0	30,668	27,917	58,585	100.0

tions of income.[3] Like many arithmetical averages it may not coincide with any one of the distributions averaged, yet it can be said to summarize them more or less adequately. Had annual distributions for fiscal years beginning February 1, instead of calendar years beginning January 1, been used, the average would probably be different, although it is less likely to be affected than any of the annual distributions. It would probably be affected appreciably by including more boom and fewer depression years, since the distribution of income apparently changes with the volume of income. An average based on data for 1927-31 would probably show more equality than a similar average based on 1929-33 data. In each case, the items (years) to be averaged are few and the changes in the volume of income from year to year appreciable. But if such an average were based on data for many years it would probably be affected only slightly by a shift from calendar to fiscal years, or by the addition or subtraction of one or two years from the period covered by the average. If it was found that the difference between this average and the individual years included in it was small, the average could be used with some confidence as a reasonable description of the distribution of annual income during the period.

Characteristics of Average Lorenz Curves

The method of averaging Lorenz curves just outlined cannot be applied generally.[4] Its value lies in the fact that if (a) each of the several annual distributions to be averaged is composed of the same receiving units, as in the sample of identical taxpayers, and, perhaps more important, (b) the receiving units are arrayed by size of income in the same order each year, the average Lorenz curve will represent the distribution of income for the series of years taken as a single accounting period. Since an accounting

[3] The weights in this case are the volumes of income represented by each curve. Other weights may be used, of course. An unweighted average (i.e., one in which each year is weighted equally) can be obtained by computing, for a series of Lorenz curves, a simple arithmetic average of the percentages of income received by each percentage of persons. As will be shown later, the weighted average has some advantage.

[4] No effort was made to ascertain the conditions necessary for its successful application. In general, it is restricted to cases where the relative distribution of income is more important than the absolute.

period of, say, two years reflects the reranking of incomes from one year to the next, the average Lorenz curve provides a bench mark against which the stability of the array of incomes during several years can be measured.

A distribution of income could be obtained for a single month or, for that matter, a single day. It would probably differ from a similar distribution for another period of like duration. Seasonal employment, the payment of dividends and interest at quarterly or semi-annual intervals, vacations without pay, the realization of profit from a sale that took several months to culminate, and a multitude of similar forces might cause considerable discrepancy between the distribution for one month and that for another. The distribution of annual income is made by adding all the income a family received at any time during the year and classifying the individual by his total income for the year. In the one case the accounting period is only one month, in the other, a year. The distribution for the year would probably not coincide with the distribution for any one of the twelve months it included. A family might be among those in the upper reaches of the distribution in one month, and among those at the very lowest end of the distribution the next; the annual income takes these shifts in rank from one month to another into account, since the family is classified by the year's total income.

The annual distribution would coincide with the average distribution for the twelve months only if each receiving unit included had the same rank each month, that is, if the persons with the highest income in January had the highest income also in February, March, etc.[5] If not, the income computed from a Lorenz curve for a specific percentage of recipients in one month would not include the same recipients as the corresponding interval the next month. In other words, this method would mean

[5] This is a more exact statement of the condition under which an average distribution of a series of short periods coincides with the distribution for a longer accounting period co-extensive with the sum of these short periods. To maintain the same rank there must be the same number of recipients. If a recipient is included in the distributions for some of the years, but not for the others, he must have $0 income throughout the entire period or the average distribution will not coincide with the distribution for the longer accounting period. The condition of identical ranking is relaxed a bit by computing the average at 5 percent intervals, since there might be some changing of ranks within the interval, but the resulting curves would be approximately the same.

adding A's income in the first month to B's income in the second month, etc., to obtain a total for the entire year, and would lead to false conclusions.

By using an annual accounting period the influence of many of the institutional arrangements that make for variation in income receipts from month to month is overcome. Semi-annual payment of interest might greatly affect the monthly distributions of income, but would have no effect on annual distributions. In some cases the effect of similar institutional arrangements, e.g., corporate dividend policy, may not work themselves out for more than a year. Numerous circumstances may influence an individual's income from one year to the next, and these, together with the institutional arrangements that have grown up in our society, may lead to considerable reranking of individual incomes from year to year.

B THE DISTRIBUTION OF INCOME FOR LONGER ACCOUNTING PERIODS

Distributions of net taxable income for two- and three-year accounting periods, and of capital gains for two- to seven-year accounting periods were prepared.[6] A comparison of the Lorenz curves for these periods with the Lorenz curves constructed by averaging the curves for the annual periods they include provides direct evidence on the extent of the reranking of income recipients and the net effect of reranking on the distribution of income.

Net Taxable Income

The results of these tabulations of net taxable income for the two- and three-year accounting periods are shown in Table 12, and graphically in the form of Lorenz curves in Chart 9. Each Lorenz curve is compared with a Lorenz curve obtained by averaging the annual distributions included in the same period on the assump-

[6] To obtain these distributions the totals for the longer periods for each individual had to be accumulated by hand from the machine tabulating tapes underlying the published income tax statistics, then tabulated. Since there were 13,000 families in the sample, this procedure was too laborious to extend to all seven years. The accumulation of capital gains was feasible, however, since 1,482 families reported only 2,573 items.

Table 12

Distribution of Net Taxable Income for 2- and 3-Year Accounting Periods

Sample of Identical Taxpayers, 1929-1931

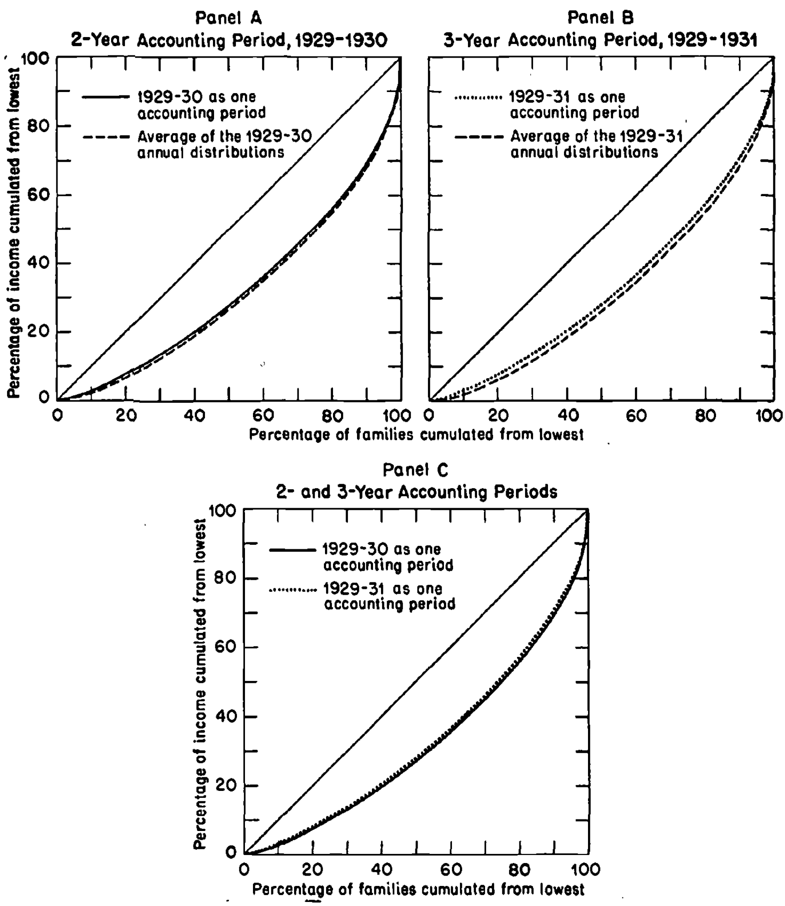
Income Group	2-Year Accounting Period, 1929-1930: Families	Net Taxable Income ($000)	Cumulated % from Lowest: Families	Cumulated % from Lowest: Income	3-Year Accounting Period, 1929-1931: Families	Net Taxable Income ($000)	Cumulated % from Lowest: Families	Cumulated % from Lowest: Income
$0	45		.3		27		.2	
1- 499	65	19	.8		39	12	.5	
500- 999	166	129	2.1	.3	84	64	1.1	.1
1,000- 1,499	402	511	5.1	1.1	172	219	2.4	.4
1,500- 1,999	1,052	1,864	13.1	4.3	262	465	4.4	.9
2,000- 2,499	1,377	3,099	23.6	9.6	590	1,340	8.9	2.6
2,500- 2,999	1,579	4,343	35.6	17.0	911	2,510	15.8	5.6
3,000- 3,499	1,684	5,472	48.3	26.3	997	3,237	23.4	9.6
3,500- 3,999	1,699	6,355	61.2	37.2	1,109	4,154	31.8	14.7
4,000- 4,499	1,289	5,446	71.0	46.5	1,152	4,894	40.5	20.7
4,500- 4,999	941	4,442	78.1	54.1	1,195	5,664	49.6	27.6
5,000- 5,999	1,077	5,858	86.3	64.1	2,183	11,978	66.2	42.3
6,000- 6,999	547	3,518	90.4	70.1	1,472	9,495	77.3	53.9
7,000- 7,999	314	2,345	92.8	74.1	868	6,457	83.9	61.8
8,000- 8,999	211	1,786	94.4	77.1	564	4,776	88.2	67.6
9,000- 9,999	147	1,391	95.5	79.5	347	3,283	90.8	71.6
10,000-14,999	345	4,136	98.2	86.6	714	8,529	96.2	82.1
15,000-19,999	113	1,930	99.0	89.8	247	4,225	98.1	87.2
20,000-29,999	74	1,757	99.6	92.8	151	3,603	99.2	91.6
30,000-39,999	23	806	99.8	94.2	47	1,591	99.6	93.6
40,000-49,999	14	622	99.9	95.3	24	1,096	99.8	94.9
50,000-74,999	9	544	99.9	96.2	11	662	99.9	95.7
75,000-99,999	3	239	99.9	96.6	9	755	99.9	96.7
100,000 & over	7	1,972	100.0	100.0	8	2,740	100.0	100.0
Total	13,183	58,585			13,183	81,748		

tion that no family was in one 5 percent interval in one year and in a different 5 percent interval the next. The Lorenz curves based on two-year and three-year income lie visibly closer (on a chart 10

CHART 9

Lorenz Curves for the Distribution of Net Taxable Income 2- and 3-Year Accounting Periods

Sample of Identical Taxpayers, 1929-1931



inches square) to the line of equal distribution than do the corresponding 'average' Lorenz curves. That this is the case indicates that there was some reranking of individuals. The 'average' curve is one limit and the line of equal distribution the other, between which the Lorenz curve for a given number of years taken as a single accounting period will lie. If there were no reranking of individuals the 'average' Lorenz curve would be identical with the Lorenz curve for the longer accounting period. But if there is reranking, the Lorenz curve for the longer period must lie nearer the line of equal distribution.[7] The limit of this reranking is the line of equal distribution which, if it described a real situation, would indicate that over a period of years all incomes averaged the same.

The comparison is between Lorenz curves that assume no reranking of recipients and those which take some reranking into account. Both are summaries of several 'annual' distributions of income, which may vary considerably. How do the curves for longer accounting periods compare with those for the annual periods included in them? The curve for the two-year accounting period 1929-30 lies nearer the line of equal distribution throughout the entire range than the curve for either 1929 or 1930. The curve for 1931, however, is nearer the line of equal distribution from about 90 to about 99 percent of the recipients included than the curve for the three-year accounting period 1929-31. This indicates the danger in drawing inferences concerning the position of a Lorenz curve for a given accounting period from the curve based on data for only a part of the period. Since no segment of an average distribution can lie outside the corresponding segment of the curve describing the least equal of any of its component an-

[7] This can be seen if we compare an average distribution for two years and a distribution for the two-year accounting period. The latter distribution is identical with the former except that two individuals, A and B, exchanged ranks in the income scale. The average distribution (in effect) attributed the income B_1+B_2 to individual B for the two years and the income A_1+A_2 to A; where $A_1<B_1$ and $A_2<B_2$. However, because of the reranking, A actually received A_1+B_2, and B received B_1+A_2.

$|(B_1+A_2)-(A_1+B_2)| < |(B_1+B_2)-(A_1+A_2)|$ since $A_1<B_1$ and $A_2<B_2$, which means that reranking reduces the difference between the incomes of A and B.

Now choose an interval on the recipient axis of the average Lorenz curve such that A has the lowest and B the highest income of all individuals included. By reducing the difference between B's and A's income, the reranking raises the minimum and reduces the maximum income included in the interval, thereby increasing (decreasing) (since the total income of all recipients in the interval is unchanged) the percentage of total income going to the recipients with incomes in the lower (higher) part of the interval, and hence pushes the section of the Lorenz curve included in the interval closer to the line of equal distribution, without affecting the other sections of the curve.

nual distributions and, since the average is one limit of a Lorenz curve for a 'long' accounting period, the Lorenz curve for a 'long' accounting period lies (for any segment) between the least equal distribution (in that segment) for any of its time components and the line of equal distribution.

Does the Lorenz curve for the three-year accounting period 1929-31 lie closer to the line of equal distribution than that for the two-year period 1929-30 because we use these particular data, or may we expect similar results whenever the accounting period is lengthened? To answer this question, two facts must be considered. First, the changes in the averages of the annual distributions, our only key to the shape and position of the Lorenz curve for the longer period, depend in part on the changes in the annual distributions. If a series of annual distributions tends to fluctuate at random within a narrow range, the addition of consecutive years will change the average less and less until it will appear to remain stationary with the addition of further years. If there is a pronounced trend in the changes from one year to another, e.g., if each distribution becomes more concentrated than the preceding, the average will follow this trend, but at a slower pace, e.g., that for 1929-32. If the trend is toward the line of equal distribution, the averages will also move toward equality, but more slowly.

Second, if there is no reranking, the average curve will be identical with the curve for the longer accounting period, and they will move together toward or away from the line of equal distribution. Or, if one of the curves included in the average lies very close to the line of equal distribution, reranking may have little effect.[8] Otherwise, the distribution of the combined incomes will lie nearer the line of equal distribution than will the average distribution.

Since the three-year total income of each person may be ob-

[8] The limiting situation is reached when one curve is the line of equal distribution. The rank of each income included in the distribution described by the line of equal distribution is $\frac{1 + \ldots + n}{n}$, and unless it were combined with another equal distribution, there would be considerable discordance between the two rankings. Nevertheless, the average and the combined distributions would be identical. For the relation of the two distributions and their average, see *Consumer Expenditures in the United States,* p. 188.

tained by adding his income in a third year to a previously determined two-year total, net taxable income may be used to answer the question. If the 1931 distribution of net taxable income is thought of as the distribution of income for a two-year accounting period, and the two-year distribution as the annual distribution for a third year, their combination would give the curve shown in Chart 9 as the curve for the three-year accounting period. Since this curve crosses the curve for 1931, the curve for one accounting period does not establish a limit for the curve for a longer accounting period that includes it.

Practically, however, it does provide a limit. The extreme case just cited, in which the additional 'year' had more than two and a half times as much income as the 'two-year' period, is not likely to occur. Three conditions must be satisfied if a curve for, e.g., a three-year accounting period is to lie further from the line of equal distribution than that for a two-year accounting period: (1) the income in the third year must be as great as or greater than that for the two preceding years combined; (2) the third year curve must lie much further from the line of equal distribution than those for the two preceding years; and (3) there must not be any substantial reranking between the two-year distribution and that for the third year.

All the data indicate that the changes in the Lorenz curve from year to year are small, and show divergent tendencies from one part of the curve to another. While income fluctuates considerably, it seldom doubles from one year to the next. As more and more years are added to the accounting period, the changes in the volume of income would have to be larger and larger for the addition of another year to move the Lorenz curve closer to the line of equal distribution as long as there is any reranking of individuals. The weight of income accumulated for several years would make peculiarities in the distribution for an additional year of negligible importance.

While we may expect the distribution for a longer accounting period to show more equality than that for a shorter, we may also expect the differences between the distributions to become smaller as the accounting period is successively lengthened. While the differences between the distributions for a two- and three-year accounting period may be substantial, the differences between those for nine- and ten-year accounting periods would probably be negligible. For example, if the income of a person who had received $90,000 during nine years approached $0 in the tenth year, he would still be classified within a narrow range on the income scale. If this analysis is correct, it would suggest the existence of an accounting period, the Lorenz curve for which would fluctuate within narrow limits and be practically the same as that for any longer period.

The length of such an accounting period would suggest the interval necessary to rid the distribution of the year to year differences introduced by institutional arrangements for the realization of income and would be of material aid in distinguishing between the differences attributable to statistical methods, i.e., use of a year as the basis of measurement and those which are more fundamental, i.e., represent differences in opportunity, ability, and the distribution of property.[9]

The extent to which a longer accounting period would give a more significant classification of individuals by income size is problematic. But, as has long been recognized, the use of the calendar year as the accounting period has led to some inequities in income taxation.[10] Moreover, consumer expenditures differ considerably at a specific income level, partly because of changes in income from the preceding year.[11] These variations would probably be reduced somewhat if the classification were based on a longer accounting period. While some variation could be expected to persist, the use of longer accounting periods would be a step toward providing a basis of income classification relatively free from the year to year fluctuations attributable to predictable institutional arrangements.

9 Many factors that make for differences are ignored here; e.g., geographic, price, and age differentials.

10 In basing the income tax upon the average of three years' income, Wisconsin tried to overcome some of the inequities of taxing fluctuating income at progressive rates. The law, faulty in that it delayed the payment of taxes too far beyond the year in which the income was received, broke down under mounting delinquency (*Report of the Tax Commission, 1934,* pp. 111-2). See also the proposals of Henry C. Simons, *Personal Income Taxation* (University of Chicago Press, 1938), p. 154, and William Vickrey, Averaging of Income for Income Tax Purposes, 47 *Journal of Political Economy,* p. 379 (June 1939).

11 *Consumer Expenditures in the United States,* pp. 160-1.

Capital Gains

Net taxable income is a universal item, i.e., one reported by every taxpayer, and every individual can be classified by the amount he receives. The fact that some persons receive $0 is in itself significant, and provides a basis for significant classification. But could capital gains of $0 during a given year be a basis for classification in the same sense? To treat all persons as recipients of capital gains, most of whom received $0, would provide a distribution of capital gains among all persons. To restrict the distribution to those who received some capital gains would provide a distribution of capital gains by their own size. Between these two poles a whole series of classifications is possible. Capital gains may be distributed among persons who received capital gains in one of two years, in two of three years, or at some time during their lives. The number of persons and of capital gains receipts will, of course, vary with the type of distribution chosen.

So long as the accounting period is fixed, this problem is not likely to lead to difficulties. Nothing is lost by moving from a distribution of capital gains by their own size to a distribution among all recipients, or among all recipients who have some other definite characteristics; if the magnitudes of these various groups are known, such a distribution can often be translated from one basis to another to suit particular purposes. When several years are treated independently, the only additional problem is one of comparability. But when an accounting period of one length is to be compared with an accounting period of another length, the question of whom to include and whom to exclude from the distribution becomes important.

Not all the capital gains realized during a year are received at one moment. Some are received in January, some in February, etc. If the accounting period is shortened to one month, fewer capital gains would probably be realized than during the calendar year that included the month. Likewise, a two-year period would probably include more persons who received capital gains than would a one-year. To compare two one-year periods with the two years as an accounting period, all who received capital gains in either year would have to be included. In each one-year period those who received no capital gains would have to be treated as having received $0 capital gains. If seven annual periods were included, all persons who received capital gains during any one of the seven years must be included in the distribution. Consequently, the Lorenz curves presented in the preceding Section must be so recast as to include persons who received capital gains at some time during the period though not necessarily in the particular year.

This type of distribution is more convenient for reading the Lorenz curves, although a distribution of capital gains among all families included in the sample of identical returns is more readily interpreted. Only 11 percent of the families included in the sample received capital gains during the seven years, and not more than 5 percent received capital gains during any one year. If the remainder are treated as having received $0 capital gains, the Lorenz curve would rise above the base line only at the 89 percent mark or beyond, depending upon the accounting period. While, theoretically, the portion of the curve above the base line would show the differences between the distributions, in practice, the curves would rise too rapidly to be readable. By omitting those who received no capital gains during the seven years—and their position can be expressed adequately by simple percentages —the Lorenz curve is a more usable device for studying differences in distributions.

The longer the period covered, the closer is the Lorenz curve to the line of equal distribution (Chart 10). Since more than one-half of the capital gains reported during the seven years 1929-35 were reported in 1929 and 1930, there is not much difference between the average Lorenz curves for three to seven years and the average curve for two years. Consequently, as the accounting period is lengthened, the spread between the average Lorenz curve and the Lorenz curve for the same number of years taken as a single accounting period widens. Is this progression toward the line of equal distribution with a lengthening of the accounting period about the same with the addition of each year, or does it diminish after two or three years have been added?

The ratios of concentration, i.e., the area enclosed by the line of equal distribution and the Lorenz curve, expressed as a percentage of the area enclosed by the line of equal distribution and the axes, provide some measure of this progression. The year to

CHART 10

Lorenz Curves for the Distribution of Capital Gains among Families Reporting Capital Gains in at Least One of the Seven Years, 1929-1935

2- to 7-Year Accounting Periods, Sample of Identical Taxpayers

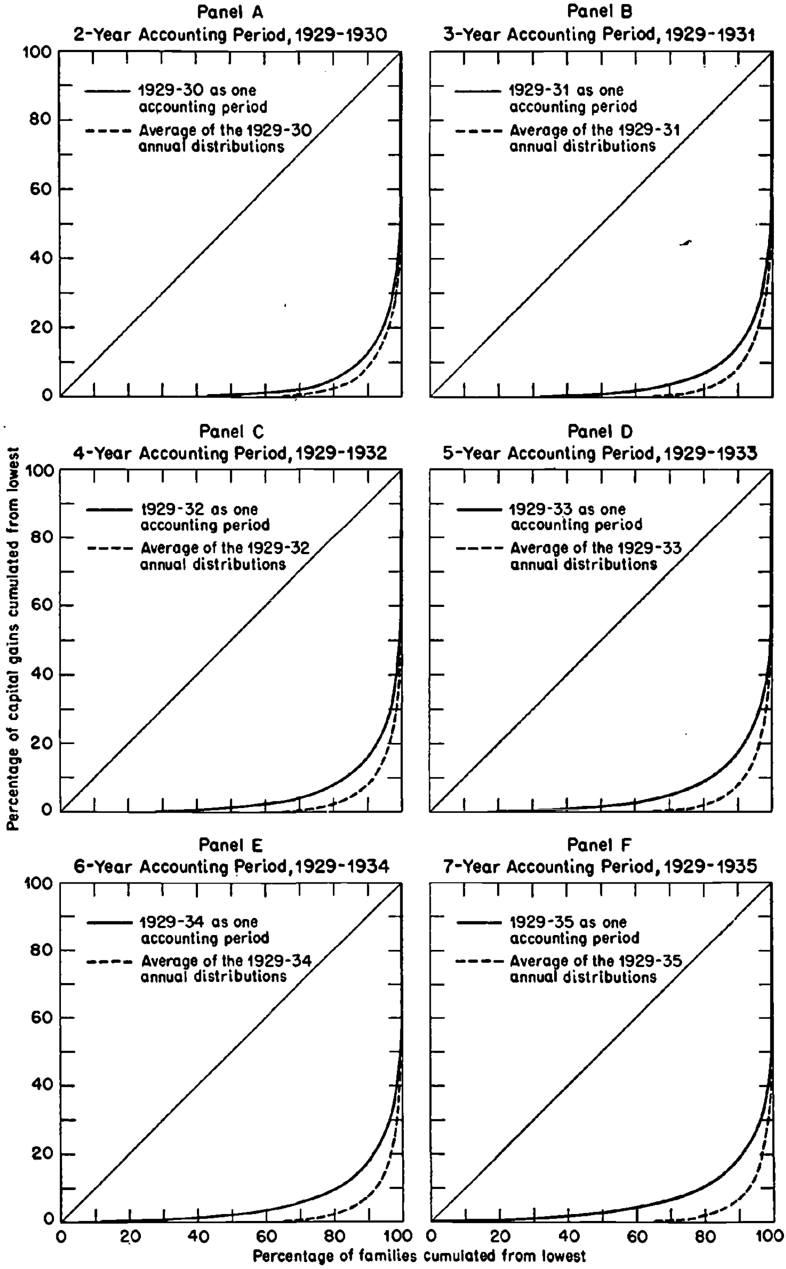
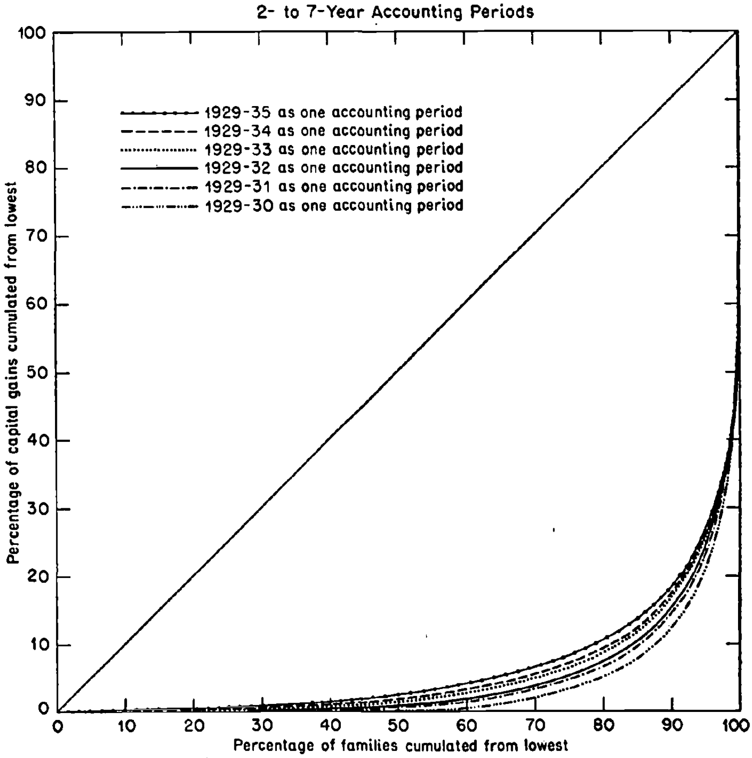


CHART 10 *(concluded)*

Panel G
2- to 7-Year Accounting Periods



year difference is .01 as another year is added to the accounting period up to 6 years and increases to .02 between the 6- and 7-year periods. Considerable differences between the 2- and 3-, and 3-

ACCOUNTING PERIOD (years)	RATIOS OF CONCENTRATION	ACCOUNTING PERIOD (years)	RATIOS OF CONCENTRATION
2	.92	5	.89
3	.91	6	.88
4	.90	7	.86

and 4-year periods, but only small differences between the remaining distributions, would have suggested that an accounting period of four years would have been sufficient to overcome all except a few of the effects of institutional arrangements on the distribution of capital gains. In the absence of a sharp break of this type, and particularly with an increase at the end of the period, we must conclude that seven years is not long enough to overcome the influence of the accounting period on the distribution of capital gains by size.

Since capital gains tend to be nonrecurring, this conclusion might have been expected. If the accounting period were lengthened to eight years, more persons would have to be included in each distribution, since some who received no capital gains during the seven years would undoubtedly have received them in the eighth year. Much of the difference in the curves is due to the progressive reduction of the number of persons who received no capital gains (Table 13). The 2-year panel of Chart 10 includes 631 persons with no capital gains; the 3-year, 477; the 4-year, 412; the 5-year, 285; and the 6-year, 167. This means that 167 persons received capital gains in 1935 who did not receive any in 1929-34. Other evidence of the irregularity of this item is that the 1,482 persons reporting capital gains reported only 2,573 items during the seven years and that at least 391 persons reported gains in only one year. No one who received a gain during the seven years was excluded in the construction of the curves and all who were included received some gains. Those who received a gain in 1929 but none in subsequent years were included in each distribution at the amount of their 1929 capital gain. Those who did not receive a gain until 1935 were included in all distributions except the 1935 at the $0 level.

While much of this apparent increase in the equality of the distribution can be attributed to the nonrecurring nature of the item, it is also true that a few persons receive capital gains every year. If the accounting period were lengthened, would the equality among these few tend to increase? As noted, reranking would yield a curve nearer the line of equal distribution than an average curve based upon the assumption of constant ranks. Whether the successive addition of time intervals to the accounting period would give a more equal distribution of capital gains among those who receive capital gains every year was not investigated.

The data do permit, however, an examination of one other question of interest. As noted in Chapter 3, the distribution of capital gains by their own size fluctuated considerably from year to year—they were more equally distributed than dividends in some years, less in others. If a longer accounting period were chosen, would the distribution of capital gains by their own size be more or less equal? Essentially this question is whether the few families who realized gains during several of the years included in the accounting period tended to realize them in amounts smaller or larger than the average. Those who received gains in only one year would not affect the distribution if the accounting period were lengthened, for their gains would be entered in a table for either accounting period at the same amount. The longer accounting period would, of course, include more capital gains, but there is no reason to think that their distribution by size would be altered appreciably. If no one received gains in two or more of the seven years, the distribution for the longer accounting period would be obtained in the same way as the combined distributions of any other two independent groups, e.g., carpenters and plumbers. Those who realized capital gains in more than one year during the longer accounting period would be entered in the distribution at the sum of their gains. If the individual gains that made up this sum were less than average, their combination might bring them closer to the mean and give a more equal distribution; it might also give entries that deviate more than the annual items from the mean, and if some of the individual items were greater than the mean, it would necessarily do so. In these cases the curve for the longer period would lie further from the line of equal distribution.

The Lorenz curves for the distribution of capital gains by their own size for 2- to 7-year accounting periods are so close together that differences between them cannot be discerned on a 10 inch square chart. Drawn on a larger scale, Lorenz curves for the top 25-30 percent of the families move away from the line of equal distribution as the accounting period is lengthened. The curves for accounting periods of various lengths cross at 70-75 percent of the families. This would indicate that the combination of gains for several years made for a slightly less equal distribution. By and large, the distribution for 1929 and 1930, which contained more than 40 percent of the items, dominated the picture.

Table 13

Cumulative Percentage Distribution of Capital Gains by 2- to 7-Year Accounting Periods
1,482 Families Reporting Capital Gains, Sample of Identical Taxpayers, 1929-1935

INCOME GROUP	TWO-YEAR No. of Families	TWO-YEAR Capital Gains	THREE-YEAR No. of Families	THREE-YEAR Capital Gains	FOUR-YEAR No. of Families	FOUR-YEAR Capital Gains	FIVE-YEAR No. of Families	FIVE-YEAR Capital Gains	SIX-YEAR No. of Families	SIX-YEAR Capital Gains	SEVEN-YEAR No. of Families	SEVEN-YEAR Capital Gains
$0	42.6		32.2		27.8		19.2		11.3			
1- 499	71.5	2.5	67.4	2.8	65.2	2.8	61.0	3.0	57.1	2.9	52.0	2.9
500- 999	80.6	5.4	78.6	6.1	77.6	6.3	74.9	6.5	72.9	6.5	68.4	6.1
1,000- 1,499	85.1	8.0	83.6	8.8	82.8	8.9	81.4	9.5	79.7	9.2	76.5	8.9
1,500- 1,999	87.8	10.1	86.6	11.0	86.0	11.2	84.7	11.6	83.6	11.4	81.2	11.2
2,000- 2,499	89.3	11.6	88.4	12.7	87.9	12.8	86.9	13.4	85.7	13.0	83.8	12.8
2,500- 2,999	90.9	13.7	90.3	15.0	89.7	14.8	88.7	15.3	87.5	14.6	86.2	14.7
3,000- 3,499	92.2	15.6	91.6	16.7	91.1	16.7	90.1	17.0	89.2	16.4	87.8	16.1
3,500- 3,999	92.8	16.7	92.2	17.6	91.8	17.7	91.0	18.2	90.2	17.6	88.8	17.1
4,000- 4,499	94.0	18.9	93.3	19.7	93.1	19.9	92.3	20.3	91.5	19.4	90.2	18.8
4,500- 4,999	94.2	19.5	93.7	20.6	93.6	20.9	93.0	21.5	92.3	20.7	91.3	20.3
5,000- 5,499	94.5	20.2	94.2	21.5	93.9	21.7	93.6	22.7	92.9	21.8	91.9	21.2
5,500- 5,999	94.9	21.3	94.6	22.5	94.3	22.5	93.9	23.2	93.3	22.5	92.4	22.0
6,000- 6,499	95.3	22.5	95.0	23.7	94.8	23.8	94.4	24.5	93.9	23.8	93.0	23.2
6,500- 6,999	95.8	24.2	95.4	24.9	95.2	24.9	94.8	25.5	94.5	25.0	93.7	24.3
7,000- 7,499	96.3	25.8	96.0	26.6	95.8	26.5	95.3	26.8	94.8	25.7	94.1	25.1
7,500- 7,999	96.7	27.2	96.3	27.5	96.1	27.6	95.8	28.1	95.3	27.0	94.6	26.3
8,000- 8,499	96.9	28.0	96.7	29.2	96.5	28.9	96.3	29.8	95.8	28.3	95.2	27.7
8,500- 8,999	97.0	28.5	96.9	29.9	96.7	29.6	96.4	30.2	95.9	28.7	95.6	28.7
9,000- 9,499	97.1	28.8	97.1	30.4	96.9	30.4	96.6	30.7	96.2	29.5	95.9	29.4

9,500- 9,999	97.2	29.5	97.2	31.0	97.0	30.7	96.7	31.2	96.2	29.7	96.0	29.8
10,000-10,499	97.4	30.4	97.4	31.9	97.2	31.8	97.1	32.4	96.6	30.9	96.4	30.8
10,500-11,999	97.7	31.8	97.6	32.8	97.6	33.3	97.3	33.5	96.9	32.1	96.7	31.8
12,000-13,999	98.1	33.8	97.9	34.7	97.9	35.1	97.8	35.8	97.5	34.4	97.2	33.5
14,000-14,999	98.4	36.0	98.3	36.8	98.3	37.0	98.1	37.2	97.9	36.6	97.6	35.1
15,000-15,999	98.5	36.5	98.3	37.2	98.3	37.0	98.1	37.2	97.9	36.6	97.7	35.7
16,000-17,999	98.7	38.2	98.5	38.2	98.4	37.9	98.4	39.3	98.2	38.1	98.0	37.3
18,000-19,999	98.9	40.5	98.7	39.8	98.6	39.5	98.5	40.2	98.3	39.0	98.2	38.4
20,000-20,999	98.9	40.5	98.7	40.4	98.7	40.6	98.7	41.3	98.5	39.9	98.4	39.5
21,000-23,999	99.0	41.2	98.9	42.4	98.9	41.8	98.7	41.8	98.5	40.4	98.4	39.5
24,000-24,999	99.1	41.9	99.0	43.1	98.9	41.8	98.7	41.8	98.5	40.4	98.4	39.5
25,000-25,999	99.1	42.8	99.1	43.8	99.0	43.3	98.9	43.1	98.7	41.5	98.5	40.0
28,000-29,999	99.1	42.8	99.1	43.8	99.0	43.3	98.9	43.1	98.7	41.5	98.5	40.0
30,000-34,999	99.1	42.8	99.1	44.7	99.1	45.0	99.1	45.5	98.9	43.7	98.8	42.6
35,000-39,999	99.3	46.2	99.3	48.0	99.3	47.0	99.1	46.4	98.9	44.5	98.9	44.0
40,000-49,999	99.5	48.9	99.5	50.5	99.5	50.5	99.3	49.7	99.2	48.3	99.1	46.5
50,000-59,999	99.5	48.9	99.5	50.5	99.5	50.5	99.4	50.9	99.2	48.3	99.2	47.6
60,000-69,999	99.5	48.9	99.5	50.5	99.5	50.5	99.4	50.9	99.3	49.7	99.3	48.7
70,000-79,999	99.6	53.7	99.6	54.9	99.6	54.6	99.5	54.7	99.4	53.0	99.4	51.6
80,000-89,999	99.6	53.7	99.6	54.9	99.6	54.6	99.5	54.7	99.5	54.9	99.4	51.6
90,000-99,999	99.7	59.5	99.7	60.5	99.7	57.2	99.7	59.5	99.5	57.0	99.4	51.6
100,000 & over	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

The number of families in each distribution is 1,482. The capital gains are:

ACCOUNTING PERIOD (years)	YEARS INCLUDED	CAPITAL GAINS	ACCOUNTING PERIOD (years)	YEARS INCLUDED	CAPITAL GAINS	ACCOUNTING PERIOD (years)	YEARS INCLUDED	CAPITAL GAINS
2	1929-30	$3,185,510	4	1929-32	$3,659,470	6	1929-34	$4,497,420
3	1929-31	3,457,620	5	1929-33	4,001,810	7	1929-35	5,268,960

C Relation of the Two- to the Three-Year Distributions of Net Taxable Income

The Lorenz curves for net taxable income for two- and three-year accounting periods reveal that income for the longer period is more equally distributed, and that the two curves vary uniformly throughout. One possible explanation is that the addition of the 1931 income of each person to his 1929-30 total made all incomes move uniformly toward the mean.[12] But there are other possible explanations, e.g., some incomes may have moved considerable distances toward the mean while others moved away compensating distances. Reflecting net effects, the Lorenz curves do not tell us anything about the character of the reranking that caused the movement toward the line of equal distribution.

Tabulation of net taxable income for the two- and three-year accounting periods gives, as a byproduct, the 1929-30 income of each person classified by his 1929-31 income. This cross-classification table provides the data necessary to gain some knowledge of the character of the reranking during this period. The means of each two-year income group are plotted, in Chart 11, against the average three-year income the members of the two-year groups received, and these points are connected by straight lines.

Also plotted is the line of regression computed for two-year incomes under $100,000. It represents the average change displayed by the incomes within this range. If each two-year income had been increased 27.1 percent and $529 added to each income, all points of the distribution below $100,000 would lie on this straight line. The incomes in each group, however, did not change in precisely the same way. The very lowest group—those who had no income in 1929-30—increased more than the average. Incomes from $1 to $3,000 and those above $20,000 increased less than the average. The remaining group, those with $3,000-20,000 incomes, however, experienced more than average increases.

In making this comparison, returns reporting more than $100,000 in 1929-30 were omitted, since the usual method of least

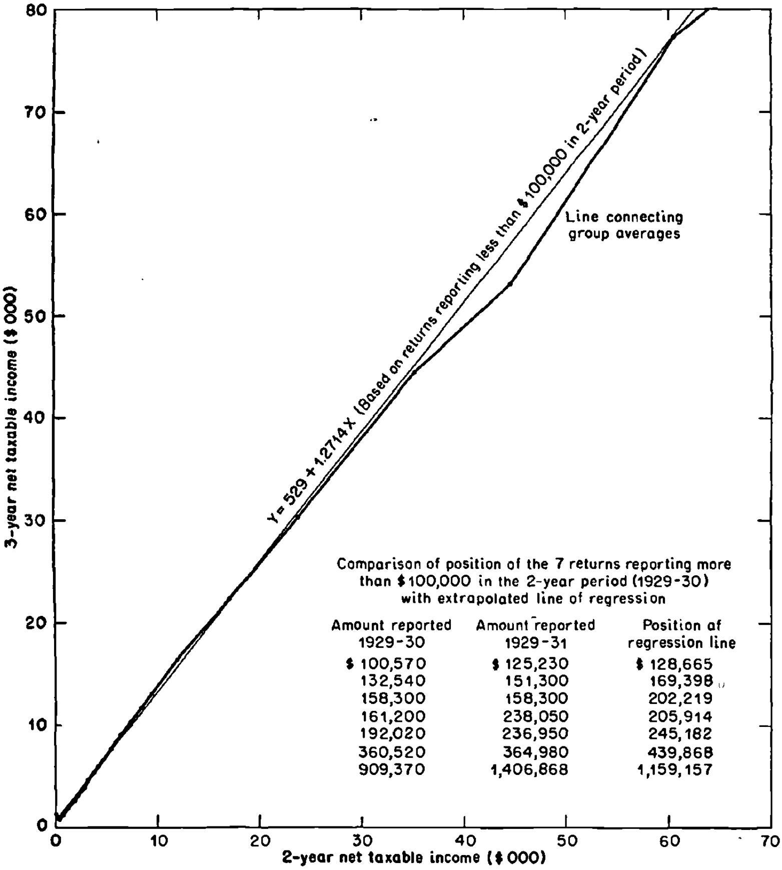
[12] For proof that shifting each income toward the mean by a given percentage of its distance from the mean will shift the Lorenz curve toward the line of equal distribution by the same percentage see *Consumer Incomes in the United States*, p. 188.

squares, used in computing the line of regression, is greatly influenced by extreme items. The one two-year income of $900,000 is sufficient to change the line of regression so much that it seems to have no relation to the remaining data; and it would re-

CHART 11

Regression of Net Taxable Income for the 3-Year Accounting Period, 1929–1931, on that for the 2-Year Accounting Period, 1929–1930

Sample of Identical Taxpayers



quire a chart 18 times as large as Chart 11 to show this extreme case on the same scale. The seven returns omitted can be compared with those distributed by extending the line of regression computed from the data for returns under $100,000. One of the two returns over $100,000, lying above the extended line of regression, is the most extreme return in the distribution.

The less than proportionate increase of the $1-3,000 group indicates that the incomes of the members lie further from the mean of the 1929-31 distribution than from the mean of the 1929-30 distribution. The more than proportionate increases of the $3,000-6,200 groups ($6,200 is the mean of the 1929-31 distribution) imply that these groups lie closer, on the average, to the mean of the three-year than to the mean of the two-year distribution, while the more than proportionate increases of the $6,200-20,000 groups indicate an increase in the distance from the mean with the addition of a year; the less than proportionate increases in the groups above $20,000 suggest the reverse. These averages suggest that reranking was confined largely to returns immediately below the mean and to the groups above $20,000, and that they were offset by opposite movements in the other groups. This, however, must remain simply a suggestion. In the first place, Chart 11 shows only the average movements of the various groups. With the exception of the very lowest group, the three-year average income is greater for each successive two-year income group. And, except for this lowest group, this chart is consistent with the assumption that there was no reranking of individuals when a year is added. All Chart 11 shows is the average income received by the individuals in each two-year group (1929-30) in 1931. These group averages cover up differences in the 1931 incomes of the individuals; and it is the reranking of individuals that determines the position and shape of the Lorenz curve.

Something of the differential treatment of individuals within each group is indicated by the coefficient of variation for the three-year net taxable income in each of the two-year groups.

These measures of dispersion (Table 14) are highest for the lowest group, fall off sharply, then fluctuate within narrow limits up to $100,000. They suggest that much of the reranking took place at the extremities of the distribution. Their behavior seems entirely reasonable. The extremes, because they are extremes, are likely to be less stable. Further, the area within which they can fluctuate is greater. High coefficients of variation, since they indicate overlapping of groups in the three-year period, are positive evidence of reranking and more confidence can be placed in them than in the regression line as a positive indication of the areas in which reranking took place.[13] Neither measure, however, can indicate the shape of the Lorenz curve or the extent to which it will differ from an average curve that assumes no reranking.

TABLE 14

Coefficient of Variation of 3-Year Net Taxable Income in Each 2-Year Group

Sample of Identical Taxpayers, 1929-1931

2-YEAR NET TAXABLE INCOME GROUP	COEFFICIENT OF VARIATION	2-YEAR NET TAXABLE INCOME GROUP	COEFFICIENT OF VARIATION
$0	4.45	$6,000- 6,999	.11
1- 499	.95	7,000- 7,999	.09
500- 999	.37	8,000- 8,999	.09
1,000-1,499	.27	9,000- 9,999	.09
1,500-1,999	.16	10,000-14,999	15
2,000-2,499	.15	15,000-19,999	.12
2,500-2,999	.14	20,000-29,999	.14
3,000-3,499	.12	30,000-39,999	.15
3,500-3,999	.10	40,000-49,999	.15
4,000-4,499	.10	50,000-74,999	.10
4,500-4,999	.10	75,000-99,999	.05
5,000-5,999	.11	100,000 & over*	1.14

* The 3-year net taxable income of each return treated separately.

Nor is the coefficient of correlation between two- and three-year incomes an entirely satisfactory measure for this purpose. True, it measures the differential movement of individual incomes, but it does not distinguish between differences that underlie reranking and differences that exist in the absence of reranking.[14] The significance of this fact may be indicated by the extreme case: a two-year income of $909,000 and a three-year income of $1,406,000. As measured by the line of regression between two- and three-year incomes, this increase is above the average. Yet there was no reranking: the same individual was in the top position in both the 1929-30 and 1931 distributions. Nevertheless, this departure from the line of regression serves to reduce the coefficient of correlation.[15] It reflects both reranking and

13 Area here refers to the 2-year groups. It is possible to speak of areas in terms of the 3-year groups, but not of areas of both 2- and 3-year groups at the same time.

14 This difficulty could, of course, be overcome by the use of rank difference correlations, but the labor involved for 13,000 individuals is prohibitive.

15 An exceptional case, such as the one cited, may produce a misleading correlation coefficient. The statement is strictly true only for normally distributed data, but has a marked effect whenever it occurs within the main body of income data.

nonlinearity because the income distribution is discontinuous. An income may shift downward a considerable distance without becoming less than the income immediately below it. Thus the coefficient of correlation indicates but does not afford conclusive proof of the reranking of individuals. Consequently, it cannot provide a clean-cut measure of the extent of reranking.

And, like the line of regression, the coefficient of correlation is greatly affected by extreme cases. One extreme income may make the coefficient approach unity despite substantial imperfections of correlation in other parts of the distribution.[16] Such coefficients are, of course, meaningless.

In the present case, for example, returns above $100,000 were omitted in computing the coefficient of correlation. The variance of the one extreme item is more than twice the variance of all other items combined. Omitting the extreme returns, the coefficient of correlation between the two- and three-year accounting periods is .98.[17]

Although the coefficient of correlation is affected by nonlinearity and by extreme cases, it does serve as a rough indicator of the extent to which incomes are reranked from one period to another. In every case the data must be examined for the presence of these two disturbing elements, and when found, care exercised in determining the significance to be attached to the resulting coefficients. A more serious limitation on their utility is that there is no way of translating them into terms of the Lorenz curve. Reranking tells us only that an average distribution based upon the assumption of no reranking errs in the direction of showing too little equality. It does not give us any clue to the shape or location of a Lorenz curve that took this reranking into account.

16 Frederick C. Mills, *Statistical Methods* (Holt, 1938, rev. ed.), pp. 370-4, gives an excellent example of this type of spurious correlation.

17 Obtained by computing the line of regression and standard deviation of each distribution, then applying the equation $r = b \div \frac{\sigma_y}{\sigma_x}$. The same approximate procedure when returns over $100,000 are included yields a correlation coefficient of .99.