This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: Changes in Income Distribution During the Great Depression

Volume Author/Editor: Horst Mendershausen

Volume Publisher: NBER

Volume ISBN: 0-870-14162-7

Volume URL: http://www.nber.org/books/mend46-1

Publication Date: 1946

Chapter Title: Changes in the Relative Position of Individual Families on the Income Scale, 1929-1933

Chapter Author: Horst Mendershausen

Chapter URL: http://www.nber.org/chapters/c5309

Chapter pages in book: (p. 81 - 113)

CHAPTER 3

Changes in the Relative Position of Individual Families on the Income Scale, 1929–1933

In the course of time, individuals and families are bound to change their relative positions on the income scale. Young persons tend to earn more as they become better qualified, and older people tend to earn less as they pass the optimum age for their occupation. The effects of structural changes and cyclical fluctuations on the various industries and occupations are different. Moreover, individual circumstances—sickness and recuperation, good and bad business fortunes—raise and lower some individual incomes relatively to others. Even Pareto, who postulated the stability of certain external aspects of income distribution, admitted that its components are continually in flux.

The income distribution for any particular year reflects all the general and special, quasi-permanent and transitory conditions that determine the incomes of individuals. From year to year as conditions change, income is redistributed. Some conditions change slowly, e.g., people's age and the structure of the economy (the relative importance of different industries, the distribution pattern of material wealth, etc.). Other conditions change more quickly, e.g., the level of business activity in a capitalist economy, and the health of individuals, or their ability to take advantage of a given economic situation. In following the incomes of the same families for a period of years, we observe the composite effect of the changes in all conditions. The longer the interval between the two terminal years, the greater the contribution of the slowly changing factors.

To measure the effect of changes in the relative magnitude of individual incomes from one year to another, the income distribution of one year is used as a base to which that of the other year is referred. Before analyzing the frequency and extent of shifts in rank and ascertaining *who* attained a higher position on the income scale and *who* a lower, we clarify our objectives.

1) We seek to measure the frequency and extent of the shifts between 1929 and 1933. It is impossible to say whether the frequency and extent of shifts observed is typical of all possible four-year intervals (see footnote 12). Supplementary data are scarce. Therefore any generalizations should be restricted at most to periods of declining business activity.

2) The same observations apply to the second question, "Who gained, who lost?" The answer refers, of course, to the groupings of the base year. If it is 1929, and if we find that the lower income recipients gained while the upper lost, this refers to the lower and upper groups of 1929. Obviously, not every man who was poor in 1929 had always been poor, or remained poor. Some had known more prosperous days, others were to improve their economic status. Clearly the rank of a particular family in the income distribution of 1929 need not be its 'normal' place. If a family had a higher income in 1933 than in 1929 it may have stepped upward on the income ladder for the first time, or it may have recovered its former position after a temporary drop.

If we knew the life incomes of all earners alive in a certain year, say 1929, i.e., the sums of their past, present, and future incomes, we could assign a 'normal' place to each family, a place unaffected by age, good or bad luck, business cycles, or structural economic change. Their relative positions in the life-income distribution could then be compared with their positions in the distribution of 1929 incomes. The extent of shifts in rank could be measured, and the groups whose relative position had improved or deteriorated established with respect to the life-income distribution. The comparison would reveal the total effect of all conditions ascribable to 1929. Then we could pass to 1933, pick out the same families, and compare their income positions with the standard. Finally, we could compare the nature and scope of the effects of particular conditions in both years and see which groups had improved their position and which had not.

Unfortunately, the restriction of our data to two years precludes the determination of life incomes, and we have refrained from computing substitutes in the form of 'business cycle, 1929–33' incomes for groups of families. Even for this limited purpose, information on two years' incomes seems too meager.[1]

Shifts in income rank were analyzed for identical samples of families in different years. In the absence of suitable data on the incomes of 'newly born' and 'deceased' families we have to content ourselves with a somewhat incomplete picture. In reality, new families continually enter the distribution, others break up, die out, and thus

[1] See Milton Friedman and Simon Kuznets, *Income from Independent Professional Practice* (National Bureau of Economic Research, 1945), and F. A. Hanna, The Accounting Period and the Distribution of Income (unpublished ms.).

vanish from the distribution. Shifts in the actual income distribution are comparable to those in a pack of cards to which new cards are added and from which old ones are dropped. Ours, however, is a constant 'pack of cards', that is to say, the changes in the 'frame' of incomes that surrounds the identical sample at any point of time are ignored. Nevertheless this picture, with all its limitations, has a good deal of interest, and its incompleteness is not too serious when the time between the terminal years is short: the shorter the interval, the fewer the 'births' and 'deaths'.

1 Frequency and Extent of Shifts in Income Rank

a PROBLEMS OF MEASUREMENT

The intensity of shifts in income rank can be defined and measured in several ways. One of the simplest is to measure the proportion of the families changing their position with respect to the median, 1929–33. The families comprise two numerically equal groups: (1) those whose 1929 incomes are below the 1929 median but whose 1933 incomes are above the 1933 median and (2) those whose 1929 incomes are above the 1929 median but whose 1933 incomes are below the 1933 median. The total in these two groups as a percentage of the number of families in the identical sample is a function of the frequency and degree of the shifts in ranks and would be at a minimum (0) if no family crossed the median line between 1929 and 1933, and at a maximum (100) if all 'submedian' families of 1929 had risen above, and all 'supramedian' families fallen below, the 1933 median.

Among various other positional values that might be considered we chose the dividing point used in the analysis of sectional inequality. In connection with that analysis it appears of considerable interest to ascertain the proportion of families that moved from the lower group in 1929 to the upper group in 1933, and conversely. Again, the families changing position consists of two numerically equal groups: (1) those whose 1929 incomes lie below the 1929 dividing point but whose 1933 incomes lie above the 1933 dividing point, and (2) those whose 1929 incomes lie above the 1929 dividing point but whose 1933 incomes lie below the 1933 dividing point. The percentages of families below the dividing point are shown for each sample in Tables 13, 16, and 17.

The proportion of families that shifted with respect to the dividing point is likewise a function of the frequency and degree of shifts in rank: assuming a minimum value (0) if all the 1929 lower group (upper group) remain in the lower group (upper group) in 1933 and a maximum value (1) if all the 1929 lower group enter the upper group in 1933, and conversely.

Both measures are simple and straightforward, but they do not cover all shifts. Families change position not only about the median, the dividing point, or any other single value, but simultaneously about many points scattered over the income scale. Neither measure records shifts that do not affect the position of a family with respect to the median and dividing point respectively. For instance, neither would register a family whose 1929 income was in the first percentile of the 1929 distribution and whose 1933 income was in the fourth percentile of the 1933 distribution. Moreover, they are insensitive to changes in the relative magnitude of incomes unaccompanied by changes in rank.

Since both measures are too crude to follow the shifts in greater detail, the coefficient of correlation between family incomes in 1929 and 1933 was used. The definition underlying this measure is best explained by calling $x_1, x_2, \ldots, x_N$ the 1929 incomes and $y_1, y_2, \ldots, y_N$ the 1933 incomes of families 1, 2, ..., N. We say there are no shifts if the differences between the income of each family and the sample mean all change in proportion to the sample standard deviation, their signs remaining constant; i.e., if for each family q (1, 2, ..., N)

$$\frac{y_q - \bar{y}}{x_q - \bar{x}} = \frac{\sigma_y}{\sigma_x}.$$

We say there are maximum shifts if the differences between the income of each family and the sample mean all change in the same proportion as the sample standard deviation, but take the opposite sign, that is, if the families change their position with respect to the mean:

$$\frac{y_q - \bar{y}}{x_q - \bar{x}} = \frac{-\sigma_y}{\sigma_x}$$

In the complete absence of shifts thus defined, the coefficient of correlation between the 1929 and 1933 family incomes will be +1, while in the case of maximum shifts it will be —1.[2] The coefficient

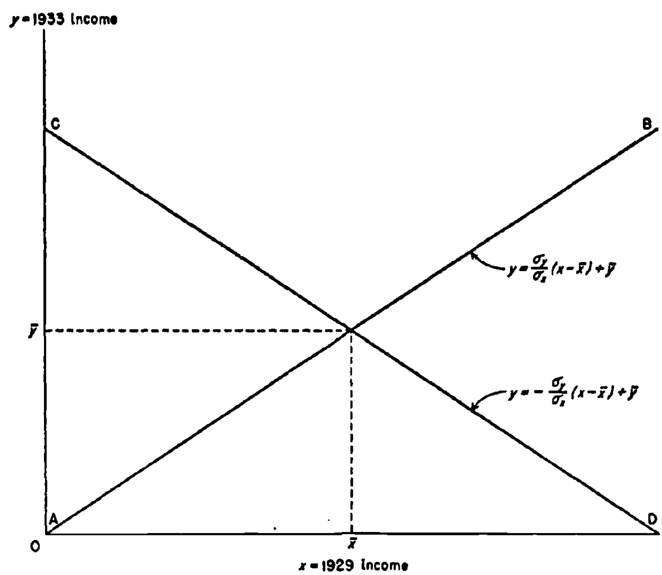
[2] In the absence of shifts we have by definition:

$y_q - \bar{y} = \frac{\sigma_y}{\sigma_x}(x_q - \bar{x})$. The coefficient of correlation is of correlation is a decreasing function of the frequency and extent of the shifts; the corresponding increasing function is 1-r.

The two extreme situations are illustrated in Chart 16. Family income in 1929 (x) is measured along the abscissa, family income in 1933 (y) along the ordinate. To each pair of incomes observed for a certain family in the identical sample a point in the x, y plane corresponds.[8] In the absence of shifts, all these points will lie on the line AB which has the slope σ_y/σ_x; in the case of maximum shifts on the line CD which has the slope $-\sigma_y/\sigma_x$. In intermediate cases the points may be scattered about a straight line of slope $b(\frac{-\sigma_y}{\sigma_x} < b < \frac{\sigma_y}{\sigma_x})$,

CHART 16

Illustration of Regression Lines



Note 2 Concluded:

$$r_{xy} = \frac{\sum_{q=1}^{N} (x_q - \bar{x})(y_q - \bar{y})}{N\sigma_x\sigma_y}.$$ Substituting, we obtain

$$r_{xy} = \frac{\frac{\sigma_y}{\sigma_x}\Sigma (x_q - \bar{x})^2}{N\sigma_x\sigma_y} = 1.$$ In the case of maximum shifts the proof proceeds in the same way and leads to $r_{xy} = -1$.

[8] In drawing the diagram it was assumed, in accordance with most of our observations in Chapters 1 and 2, that $\bar{x} > \bar{y}$, $\bar{v}_x' < \bar{v}_y'$.

or may follow (be scattered about) some sort of curve.[4] In computing the coefficient of correlation, we establish the ratio between the slope of the straight line, fitted to the observations by least squares, and the slope of the line *AB*.[5]

Another aspect of shifts in income rank merits separate attention. As we leave the base year, the individuals in certain income groups tend to scatter. In a later year the members of a certain base-year group are dispersed among many groups. When the dispersion varies in degree from one base-year group to another we speak of differences in homogeneity—or heterogeneity—of the various base-year groups with respect to income change. A group of families is perfectly homogeneous in this sense if its members move *en bloc* to another income group as time goes on—or remain *en bloc* in the old group—and the more they scatter the more heterogeneous is the group.

Heterogeneity with respect to income change can be measured by the standard deviation of the income distribution formed by the members of a base-year group in the other year. For instance, the heterogeneity of the 1929 income group $1,000-1,499 can be measured by the standard deviation of the distribution of its members among the various 1933 income groups.

We determined these standard deviations by the same procedure as we measure income dispersion in general: the families in a certain income group in a given year are assumed to have the same, i.e., the mean, income of that group. This procedure underestimates the standard deviations of individual incomes because it neglects all intragroup variation. However, there is not necessarily any systematic variation in the *degree* of bias within the various income groups.

These standard deviations measure dispersion of incomes in terms of dollars. For certain purposes it may be of interest to compare the measures of absolute dispersion with the income level of

[4] No straight line fitted by the method of least squares can have a slope outside the indicated limits since the coefficient of correlation cannot assume values outside the limits +1, −1.

[5] The slope of the fitted line obtained by minimizing deviations in the *y* direction is

$$b = \frac{\sigma_y}{\sigma_x} r_{xy}$$

The slope of the *AB* line is $b' = \sigma_y/\sigma_x$. Therefore $b/b' = r_{xy}$. If the line is fitted by minimizing in the *x*-direction the resulting slope must be compared with $1/b'$ and the value of the ratio is the same.

the group. It may be argued that a certain degree of 'absolute' heterogeneity with respect to income change, as expressed by the standard deviation, is more significant the smaller the mean income of the group in the other year. For example, if the incomes of the members of the 1929 income group $1,000-1,499 have a standard deviation of $500 in 1933, heterogeneity with respect to income change may be taken as more significant if their mean 1933 income is $400 than if it is $800. 'Relative' heterogeneity may be measured by dividing the standard deviations by the corresponding mean incomes. In our example this amounts to dividing $500 by $400, or by $800. Thus we obtain the coefficient of variation of the 1933 incomes received by members of the 1929 group $1,000-1,499.

Obviously, it is unnecessary to choose the earlier year as the base; the analysis may just as well be based on the later year. Each procedure answers a specific question. If 1929 is the base year, we ask: to what extent are the 1933 incomes of families with (practically) identical 1929 incomes spread out over the income scale? If 1933 is the base year, we ask: to what extent are the 1929 incomes of families with (practically) identical 1933 incomes spread out? In the first, we observe the end result of a process of dispersion, in the second, the starting point of a process of convergence. Both years were used, since both questions are relevant and the answer to one does not automatically give the answer to the other.

b SHIFTS ABOUT POSITIONAL VALUES AND CORRELATION OF 1929 AND 1933 FAMILY INCOME

The percentage of families changing their position with respect to the median of the income distribution ranges from 15 (Richmond) to 32 (Butte); in 29 of the 33 cities from 20 to 30 per cent (Table 28). During the four years, 70 to 80 per cent of the families in the various city samples retained their positions with respect to the median income—either below or above. Twenty to 30 per cent shifted, half rising from below the 1929 median to above the 1933 median and the other half dropping from above the 1929 median to below the 1933 median.

Results are similar for the two tenure groups, but in most (27) cities shifts in the income rank of tenants were somewhat greater than in those of owners. Apparently the relative income status of owners is more stable than that of tenants.[6]

[6] The median incomes of the various samples are shown in Table 29.

TABLE 28

Percentage of Families Shifting Position with Respect to the Median, 1929–1933

Identical Samples: Entire-city, Tenant, Owner-occupant

	ENTIRE-CITY (1)	TENANT (2)	OWNER-OCCUPANT (3)
Atlanta	17.6	19.7	20.2
Birmingham	23.3	23.9	24.2
Boise	23.1	23.9	22.3
Butte	32.1	32.8	31.0
Cleveland	26.2	26.1	25.5
Dallas	25.2	25.4	23.4
Des Moines	24.3	24.6	23.6
Erie	26.0	26.3	26.6
Indianapolis	22.8	23.0	22.0
Lansing	28.7	29.5	27.2
Lincoln	22.4	24.0	22.0
Little Rock	21.3	22.9	20.8
Minneapolis	25.4	25.1	24.6
Oklahoma City	27.6	27.9	25.2
Peoria	26.2	28.0	24.6
Portland, Me.	24.6	25.6	20.7
Portland, Ore.	26.8	29.0	25.1
Providence	24.5	27.4	24.2
Racine	31.6	31.8	31.2
Richmond	15.3	15.0	16.3
Sacramento	24.5	26.5	20.5
St. Joseph, Mo.	20.7	20.9	22.0
St. Paul	23.7	27.9	23.5
Salt Lake City	25.4	28.0	24.8
San Diego	26.4	27.3	24.8
Seattle	26.7	27.3	25.9
Springfield, Mo.	25.6	32.3	14.3
Syracuse	25.9	25.2	27.0
Topeka	21.8	24.4	21.5
Trenton	26.5	28.6	24.2
Wheeling	25.7	26.8	24.0
Wichita	25.9	27.4	24.4
Worcester	25.6	27.2	19.7

The coefficient of correlation, a measure of the *weakness* of shifts in income rank, of 1929 and 1933 income ranges from .56 (Racine) to .89 (Richmond) (Table 30). Thus, according to this measure shifts were greatest in Racine, least in Richmond.[7]

[7] Since the percentage of shifts about the median and the correlation coefficient are quite different measures of changes in income rank, it is not surprising that the ranking of the cities according to one measure differs from the ranking according to the other. However, the actual differences are small.

TABLE 29

Median Incomes

Identical Samples: Entire-city, Tenant, Owner-occupant

	ENTIRE-CITY		TENANT		OWNER-OCCUPANT	
	1929	*1933*	*1929*	*1933*	*1929*	*1933*
	(1)	(2)	(3)	(4)	(5)	(6)
			(dollars)			
Atlanta	1,264	790	990	637	2,102	1,446
Birmingham	1,262	560	1,083	468	1,829	927
Boise	1,487	1,083	1,440	1,069	1,564	1,101
Butte	1,715	725	1,674	703	1,795	765
Cleveland	1,639	922	1,544	897	1,820	969
Dallas	1,716	1,202	1,592	1,112	2,022	1,413
Des Moines	1,557	1,110	1,488	1,060	1,670	1,177
Erie	1,506	748	1,446	744	1,609	754
Indianapolis	1,760	1,141	1,598	1,045	2,141	1,353
Lansing	1,683	878	1,536	815	1,862	979
Lincoln	1,513	1,024	1,429	996	1,668	1,065
Little Rock	1,441	840	1,239	698	1,933	1,194
Minneapolis	1,650	1,136	1,526	1,064	1,857	1,248
Oklahoma City	1,666	1,069	1,546	1,006	1,945	1,221
Peoria	1,580	1,064	1,472	992	1,722	1,163
Portland, Me.	1,653	1,251	1,566	1,190	1,945	1,427
Portland, Ore.	1,458	833	1,320	730	1,671	995
Providence	1,509	1,038	1,401	952	1,784	1,222
Racine	1,618	656	1,575	634	1,682	675
Richmond	1,414	1,132	1,134	944	2,047	1,626
Sacramento	1,797	1,309	1,684	1,242	1,997	1,443
St. Joseph, Mo.	1,432	1,086	1,338	1,014	1,632	1,198
St. Paul	1,484	1,066	1,356	927	1,686	1,248
Salt Lake City	1,591	983	1,464	893	1,784	1,127
San Diego	1,556	1,142	1,492	1,116	1,676	1,186
Seattle	1,638	960	1,538	918	1,774	1,032
Springfield, Mo.	1,310	847	1,212	758	1,451	971
Syracuse	1,672	1,018	1,528	925	1,921	1,173
Topeka	1,462	974	1,352	891	1,621	1,093
Trenton	1,323	840	1,208	737	1,474	958
Wheeling	1,302	790	1,289	800	1,326	775
Wichita	1,536	906	1,475	871	1,648	974
Worcester	1,656	1,185	1,535	1,100	1,996	1,424

Again in most (29) cities tenant incomes shifted more. The correlation between 1929 and 1933 incomes tends to be lower for tenant than for owner families.

On the whole, tenant incomes show smaller declines than owner incomes during the depression (see Ch. 1) and are subject to greater shifts in rank. But is there a general inverse correlation between the rate of income decline and the frequency and degree of the shifts

TABLE 30

Coefficient of Correlation between Family Incomes in 1929 and 1933
Identical Samples: Entire-city, Tenant, Owner-occupant

	ENTIRE-CITY (1)	TENANT (2)	OWNER-OCCUPANT (3)
Atlanta	.767	.762	.735
Birmingham	.728	.727	.697
Boise	.753	.713	.781
Butte	.637	.586	.685
Cleveland	.761	.768	.755
Dallas	.732	.705	.741
Des Moines	.790	.705	.839
Erie	.745	.739	.739
Indianapolis	.801	.751	.836
Lansing	.707	.644	.749
Lincoln	.771	.736	.781
Little Rock	.776	.732	.815
Minneapolis	.716	.666	.741
Oklahoma City	.665	.550	.716
Peoria	.762	.688	.778
Portland, Me.	.766	.751	.772
Portland, Ore.	.692	.616	.720
Providence	.738	.718	.727
Racine	.700	.662	.681
Richmond	.856	.892	.813
Sacramento	.742	.683	.759
St. Joseph	.856	.840	.865
St. Paul	.761	.652	.797
Salt Lake City	.747	.707	.765
San Diego	.633	.629	.635
Seattle	.729	.677	.747
Springfield, Mo.	.753	.741	.765
Syracuse	.754	.686	.763
Topeka	.773	.711	.794
Trenton	.728	.654	.768
Wheeling	.745	.737	.742
Wichita	.669	.615	.704
Worcester	.811	.782	.808

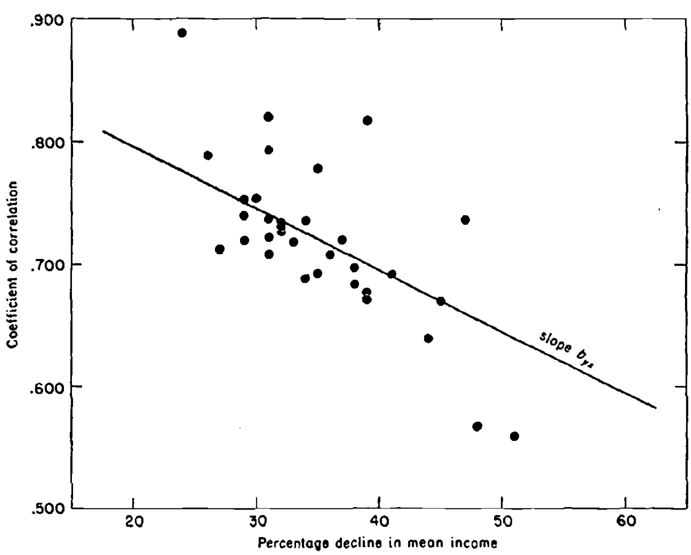
In computing the coefficients for tenants and owners in Atlanta, Butte, Minneapolis, Providence, St. Paul, and San Diego, we used the estimate of 1933 mean incomes of the various 1929 classes that are based on the assumption of equal cell means within each 1933 income class (see Sec. 3, where this method is described for the entire-city analysis of intraclass variance). Since these estimates do not differ much from the known correct means, a recomputation of these coefficients seemed unnecessary.

as measured by 1-*r*? If we take each of the 33 entire-city samples as one observation, and correlate the city's rate of income decline with its correlation coefficient of 1929 and 1933 incomes, $r = -.55$. The higher the rate of income decline, the weaker the correlation of

1929 and 1933 incomes, that is, the more extensive the shifts in income rank. Obviously, the relation between the two aspects of income change is far from close; nevertheless, a tendency in the direction indicated is obvious (see Chart 17).[8] As far as the various cities are concerned, shifts in income rank are wider the greater the impact of the depression on the city's income level. Yet for the two tenure groups, the inverse relation is found to hold: shifts in rank are more marked among the incomes of tenants, whose mean income declines less.

CHART 17

Percentage Decline in Mean Incomes and Coefficient of Correlation 33 Entire-city Samples



These two contradictory trends can be ascribed to the divergent relations between income level and rate of income decline observed for the various cities and for the two tenure groups. The mean income of the tenure group with the higher income level (owners) declines more during the depression; but cities with a relatively

[8] If the rate of decline of the mean income is replaced, in the correlation, by the dollar difference between 1929 and 1933 mean incomes, the negative correlation is slightly higher (—.62). With 33 observations, both —.55 and —.62 must be considered statistically significant on the 5 per cent level of probability.

high income level tend to show relatively small declines in income.[9] If we correlate income level and the frequency and degree of shifts in income rank from 1929 to 1933, the correlation is inverse, whether the observations are for different cities or different tenure groups. The mean income of owners exceeds that of tenants, and the income ranks of owners shift less; the higher a city's 1933 income level the weaker do shifts in its family income ranks tend to be.[10] Thus, shifts apparently tend to be consistently greater among the poorer groups and less among the wealthier.

Returning to the measurement of the frequency and extent of the shifts in income rank, we determine what percentage of all families change their positions with respect to the dividing point used in the analysis of sectional inequality. What percentage leave the lower group and enter the higher?

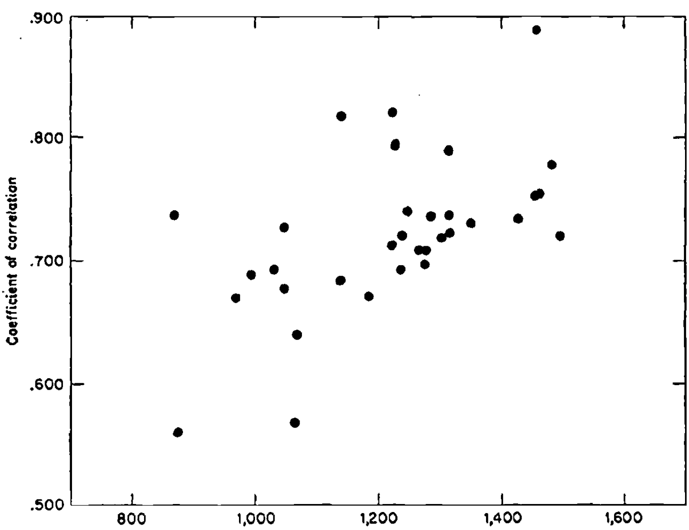
In the various entire-city samples, more than 13 but less than 26 per cent of the families are involved in transfers between the lower

[9] See Ch. 1. The relation for the various cities was found to be open to doubt from a statistical point of view.

[10] The correlation between 1933 mean income and the correlation coefficient of 1929 and 1933 income is $r = .45$ (see Chart 18). If 1929 mean income is used instead of 1933 mean income we obtain an insignificant positive correlation, $r = .12$.

CHART 18

Coefficient of Correlation and Average Income, 1933
33 Entire-city Samples



and the higher groups (Table 31). In the samples of the two tenure groups the minimum percentage of shifts is 11, the maximum 28. The great majority of families maintain their positions in the lower or upper group during the four years. Again, the samples in Richmond and Atlanta show the least shifting, while Racine and Butte are among the samples with the greatest. Comparison of corresponding entries in Tables 31 and 28 reveals that the percentage of families shifting positions about the dividing point is always smaller

TABLE 31

Percentage of Families Shifting Position with Respect to the Dividing Point, 1929–1933

Identical Samples: Entire-city, Tenant, Owner-occupant

	ENTIRE-CITY (1)	TENANT (2)	OWNER-OCCUPANT (3)
Atlanta	14.2	11.4	19.8
Birmingham	18.8	16.6	24.0
Boise	18.8	19.8	16.4
Butte	24.4	22.4	26.4
Cleveland	23.2	21.6	24.2
Dallas	21.6	20.8	23.4
Des Moines	18.8	19.0	20.0
Erie	22.4	20.2	24.2
Indianapolis	20.8	20.0	21.8
Lansing	25.8	24.4	26.0
Lincoln	18.0	17.0	19.0
Little Rock	19.0	17.8	20.6
Minneapolis	20.2	18.8	22.6
Oklahoma City	24.6	24.0	28.4
Peoria	20.8	19.6	22.0
Portland, Me.	19.0	19.0	19.6
Portland, Ore.	20.2	19.2	22.2
Providence	18.8	16.8	21.6
Racine	25.8	24.6	26.6
Richmond	14.0	13.0	16.0
Sacramento	20.8	21.4	20.6
St. Joseph	16.0	14.6	17.6
St. Paul	16.4	14.8	18.4
Salt Lake City	21.6	20.2	23.0
San Diego	21.4	22.0	21.0
Seattle	25.2	24.6	25.0
Springfield, Mo.	15.4	13.4	17.8
Syracuse	23.0	20.2	26.2
Topeka	16.8	16.2	18.2
Trenton	14.8	12.6	17.4
Wheeling	18.2	17.2	19.2
Wichita	22.2	23.0	22.0
Worcester	19.2	19.2	19.6

than the percentage of shifts about the median. This is to be explained on technical grounds.[11]

The conclusions reached in the preceding chapter must be understood in the light of these findings. It was observed that income inequality between the lower and upper groups increased from 1929 to 1933. We now see that the effects, upon the *individual* family, of the increasing relative spread between the mean incomes of the two groups are somewhat mitigated by the fact that between 1929 and 1933 some of the lower entered the upper group and vice versa. In the various cities 6 to 13 per cent of families went one way and the same percentage the other. The great majority, however, did not shift between the two groups.

In attempting to appraise the results, we must keep in mind that they refer to the four years 1929–33. Both the length and the historical character of this period affect the thoroughness with which income ranks change. To the extent that the shifts are attributable to the differential effects of both structural change in the economic system and advancing age on the incomes of individual families, they must be expected to increase with the length of the interval. Cyclical variations are less likely to lead to a uniform increase in the frequency and extent of the shifts as time passes; nor will 'chance' events have such an effect, unless they are cumulative in one direction.

Coefficients of correlation computed on the basis of identical samples of Wisconsin state income tax returns and supplied to the

[11] It is not surprising that the percentage of shifts about the dividing point always falls short of the percentage of shifts about the median. The maximum for the second set of percentages is lower than the maximum for the first set. The highest percentage of families that can possibly shift their position with respect to the median is 100; on the other hand, the maximum shifting about the dividing point is twice the percentage of families in the smaller of the two groups created by dividing the income distribution.

Suppose the dividing point in both years is located in such a way as to place the lower 65 per cent of income recipients in the lower group, the upper 35 per cent in the upper group. Then the maximum percentage of total families that may shift about the dividing point is 70. In that case the entire upper group would enter the lower group and be replaced by an equal number of former members of the lower group. If there is no particular 'bunching' of shifts, e.g., an unusually intense shifting of supramedian incomes about the dividing point, we must expect the percentage of shifts about the median to exceed the percentage of shifts about the dividing point.

Since the percentage of families in the smaller—usually the upper group—is lower among tenants than among owners (see Tables 16 and 17) the maximum percentage of shifts about the dividing point is smaller for tenants. This is the reason why tenants show fewer shifts about the dividing point in almost all cities (see Table 31). Table 31 does not contradict the finding that tenant incomes shift more than owner incomes.

author by Frank A. Hanna cast some light on the effect of cyclical variations on the frequency and extent of the shifts in income rank during 1929–35. The coefficients are:

$$r_{30,29} = .86 \qquad r_{32,29} = .70 \qquad r_{34,29} = .65$$
$$r_{31,29} = .78 \qquad r_{33,29} = .64 \qquad r_{35,29} = .69$$

Until 1933, these coefficients shrink as we depart from the base year 1929 both in time and in degree of economic prosperity; but they tend to rise during the recovery beginning in 1933.[12] This seems to indicate that as business activity returns to the level of the base year, individual incomes *tend* to regroup themselves in the old way. One might indeed advance the hypothesis that cyclical variations in economic activity produce cyclical variation in the frequency and extent of the shifts, and that the cycles of the frequency and extent of the shifts are superimposed upon a trend-like increase, which in turn is ascribable to aging of the population and structural economic change.

C HETEROGENEITY OF GROUPS WITH RESPECT TO INCOME CHANGE[13]

Let us consider first the heterogeneity of the 1929 income groups. The standard deviations in Table 32 measure the dispersion—in dollars—of the 1933 incomes of families that were in the same 1929 income group. The measures, given for each 1929 income group separately, are supplemented by the standard deviation of all 1933 incomes (last column). Between 1929 and 1933, all 1929 income groups are broken up and their members scattered over several income brackets. The degree of dispersion shows a characteristic variation from group to group, in the several cities.

Absolute heterogeneity with respect to income change is relatively high among families reporting 'no income' in 1929, although their 1933 incomes show in general a smaller standard deviation than the average for all families. As we go up the income scale stand-

[12] The coefficients of correlation are computed from a sample of 13,184 identical taxpayers who filed Wisconsin state income tax returns for each year 1929–35.

It will be noted that the correlation coefficient for 1929–33 is lower than most of the coefficients in Table 30. That is to say, the majority of the Financial Survey samples show somewhat less shifting of income rank than the Wisconsin tax samples. See also the coefficients computed for identical samples of professional workers in Friedman and Kuznets, *op. cit.*, Table 56.

[13] This analysis, confined to the entire-city samples, is based on the cross-classification tables in Appendix B. For the procedure followed in estimating the 1929 and 1933 mean incomes of the cells, see Appendix A 2.

TABLE 32: Standard Deviations of the 1933 Incomes of Families for Each 1929 Income Group

Identical Samples: Entire-city

	1929 INCOME GROUP (DOLLARS)											
	No income	*1–249*	*250–499*	*500–749*	*750–999*	*1,000–1,499*	*1,500–1,999*	*2,000–2,999*	*3,000–4,499*	*4,500–7,499*	*7,500 & over*	*All families*
Atlanta	1,202	205	282	309	374	549	589	850	1,084	1,937	4,008	1,428
Birmingham	638	233	246	375	349	508	561	769	960	1,486	2,699	955
Boise	639	350	294	351	431	435	577	730	986	1,785	2,935	993
Butte	513	359	577	655	445	508	697	895	1,292	1,545	4,916	1,316
Cleveland	731	724	384	417	537	533	650	820	1,067	1,865	4,638	1,460
Dallas	1,576	345	365	350	378	472	615	797	1,045	2,028	4,603	1,329
Des Moines	1,085	319	335	385	396	547	683	730	1,058	1,725	4,889	1,325
Erie	638	191	310	438	368	459	582	826	908	1,890	3,446	1,069
Indianapolis	692	199	290	419	759	698	663	783	1,036	1,889	3,823	1,548
Lansing	590	250	506	323	387	444	496	704	984	1,198	4,187	1,035
Lincoln	682	293	516	342	466	692	594	799	963	1,592	3,334	1,265
Little Rock	517	178	212	276	391	614	577	771	977	1,550	3,430	1,221
Minneapolis	538	417	268	342	387	480	593	786	1,029	1,875	3,692	1,234
Oklahoma City	1,135	256	658	504	442	560	585	905	1,170	1,968	3,894	1,286
Peoria	638	323	373	360	380	481	576	766	1,075	1,796	3,973	1,314
Portland, Me.	740	433	224	452	436	769	540	738	1,026	1,743	4,052	1,399
Portland, Ore.	633	348	304	538	367	494	618	802	1,006	1,656	3,698	1,046
Providence	597	382	309	315	564	477	739	949	1,071	1,953	4,271	1,443
Racine	593	269	327	602	600	445	560	716	949	1,757	2,559	919
Richmond	551	240	185	266	304	461	503	753	891	1,622	3,610	1,452
Sacramento	1,769	401	1,297	387	408	584	598	788	1,033	1,733	4,240	1,325
St. Joseph	417	201	272	402	356	428	562	767	857	2,158	3,536	1,316
St. Paul	825	240	279	493	380	440	553	726	952	1,976	4,956	1,203
Salt Lake City	691	310	425	300	380	548	652	816	1,039	1,793	3,398	1,204
San Diego	777	378	333	540	377	514	573	744	1,128	1,674	3,236	1,042
Seattle	811	340	318	343	569	597	583	773	1,071	1,653	3,090	1,073
Springfield, Mo.	393	164	258	354	365	440	569	770	1,314	1,733	4,031	1,041
Syracuse	388	358	346	283	324	423	558	752	1,057	1,550	6,277	1,384
Topeka	608	263	257	279	394	421	534	729	1,030	1,984	4,466	1,197
Trenton	435	352	315	344	389	605	550	712	994	1,466	2,779	941
Wheeling	827	283	335	432	426	458	561	743	926	2,253	3,419	976
Wichita	1,353	250	535	281	365	618	528	797	1,298	1,692	3,305	1,095
Worcester	1,141	739	409	348	398	461	559	940	1,032	1,856	3,599	1,531

ard deviations decline, reaching a minimum in one of the three next-to-lowest groups: these groups show the least heterogeneity in dollar terms. Thereafter, absolute heterogeneity increases uniformly, reaching a maximum in the top group. The standard deviations of the two—and frequently three—highest groups exceed the corresponding standard deviations for all families.

Heterogeneity in relative terms, i.e., the standard deviations of Table 32 related to the corresponding means of 1933 incomes, varies less from group to group (Table 33). In general, relative heterogeneity is highest in the 'no income' group. As we go up the income scale, relative heterogeneity declines and reaches a minimum, usually in one of the three groups between $1,500 and $4,499. The decline is rather irregular, and a secondary maximum can frequently be distinguished between $500 and $1,499. In the two highest groups all coefficients of variation rise, though not markedly. In most (28) cities, they show another secondary maximum in the top group.

The group-wise measures show values smaller than the coefficients of variation for all 1933 incomes, except in the 'no income' group and, in many instances, the next-to-lowest group. None of the secondary maxima reaches the values of the 'over-all' coefficients.

The following broad conclusions can be drawn from these observations. Heterogeneity with respect to income change tends to be most pronounced in the extreme income groups. If measurement is in dollars, the top income groups lead, followed by the lowest, and finally by the middle. Heterogeneity in relative terms is highest for the income groups at the bottom of the scale; then follow the top or, in some cases, one of the central groups, and finally the other middle groups.

The relatively high heterogeneity of the extreme groups with respect to income change indicates that the families in them tend to scatter more widely over the income scale during the four years than do the families in the central groups. In this sense, shifts in income rank affect the extreme groups most. The idea of certain income groups rising or falling *en bloc,* often useful as a rough approximation to reality, is least satisfactory when we are concerned with very low or high incomes.

We pass now to the analysis of the 1933 income groups. The standard deviations (Table 34) measuring the absolute dispersion, and the coefficients of variation (Table 35) measuring the relative

TABLE 33: Coefficients of Variation of the 1933 Incomes of Families for Each 1929 Income Group

Identical Samples: Entire-city

	1929 INCOME GROUP (DOLLARS)											
	No income	*1–249*	*250–499*	*500–749*	*750–999*	*1,000–1,499*	*1,500–1,999*	*2,000–2,999*	*3,000–4,499*	*4,500–7,499*	*7,500 & over*	*All families*
Atlanta	2.473	.869	.842	.690	.617	.632	.472	.485	.457	.544	.592	1.168
Birmingham	1.743	1.084	.872	1.019	.797	.767	.590	.573	.510	.591	.705	1.099
Boise	1.206	1.036	.770	.652	.650	.480	.458	.424	.443	.512	.495	.813
Butte	1.386	1.442	1.414	1.305	.859	.769	.882	.664	.644	.613	.805	1.239
Cleveland	1.929	2.117	.912	.827	.843	.685	.624	.565	.534	.606	.716	1.180
Dallas	1.854	1.162	.835	.624	.517	.498	.488	.471	.466	.668	.690	.931
Des Moines	2.633	1.127	.848	.724	.593	.600	.549	.439	.459	.544	.637	1.009
Erie	2.573	.838	.981	1.009	.754	.700	.613	.627	.482	.624	.642	1.103
Indianapolis	2.149	.930	.775	.871	1.060	.780	.569	.470	.448	.549	.657	1.045
Lansing	1.359	.847	1.150	.641	.660	.609	.522	.539	.519	.488	.765	.969
Lincoln	1.409	.885	1.112	.714	.711	.799	.506	.487	.411	.476	.563	.984
Little Rock	1.901	.978	.768	.667	.622	.711	.509	.496	.460	.558	.603	1.071
Minneapolis	1.868	1.171	.608	.626	.543	.530	.489	.477	.455	.539	.604	.914
Oklahoma City	1.779	1.016	1.552	.898	.656	.629	.511	.585	.559	.727	.769	1.009
Peoria	1.415	1.009	.856	.641	.538	.530	.496	.485	.482	.521	.617	1.010
Portland, Me.	1.893	1.136	.557	.773	.602	.774	.420	.425	.426	.497	.602	.958
Portland, Ore.	1.670	1.265	.869	1.121	.634	.636	.584	.557	.493	.570	.693	.999
Providence	1.983	1.322	.655	.567	.780	.543	.616	.574	.466	.555	.607	1.097
Racine	1.836	1.007	.985	1.326	1.113	.754	.693	.617	.598	.735	.580	1.051
Richmond	2.103	.805	.476	.512	.437	.453	.353	.388	.331	.400	.547	.996
Sacramento	3.567	1.162	2.411	.650	.544	.596	.475	.443	.429	.524	.688	.886
St. Joseph	1.931	.731	.687	.728	.472	.451	.427	.437	.365	.535	.520	1.002
St. Paul	1.960	1.039	.740	.913	.596	.503	.447	.418	.380	.573	.594	.964
Salt Lake City	1.763	1.051	.986	.617	.589	.633	.574	.525	.466	.572	.613	.973
San Diego	1.914	1.277	.753	.887	.482	.506	.450	.455	.511	.556	.719	.816
Seattle	1.273	1.086	.731	.636	.826	.705	.524	.530	.537	.591	.640	.906
Springfield, Mo.	1.836	.739	.727	.653	.570	.526	.493	.489	.534	.547	.719	.993
Syracuse	1.610	1.155	.742	.575	.518	.516	.493	.493	.475	.488	.802	1.092
Topeka	1.404	1.056	.689	.559	.570	.470	.444	.426	.443	.585	.645	.976
Trenton	2.458	1.253	.788	.667	.631	.757	.506	.451	.444	.455	.596	.914
Wheeling	2.440	1.040	.833	.845	.668	.551	.502	.508	.447	.678	.607	.985
Wichita	2.010	.853	1.241	.601	.603	.747	.482	.541	.633	.601	.670	.961
Worcester	2.547	1.661	.861	.589	.523	.488	.467	.550	.427	.506	.480	1.052

TABLE 34: Standard Deviations of the 1929 Incomes of Families for Each 1933 Income Group

Identical Samples: Entire-city

	1933 INCOME GROUP (DOLLARS)											
	No income	*1–249*	*250–499*	*500–749*	*750–999*	*1,000–1,499*	*1,500–1,999*	*2,000–2,999*	*3,000–4,499*	*4,500–7,499*	*7,500 & over*	*All families*
Atlanta	1,828	715	774	913	794	1,146	1,147	1,204	2,047	3,786	3,522	1,941
Birmingham	1,564	741	997	1,035	1,253	1,176	1,418	1,752	3,028	4,077	6,611	1,750
Boise	989	651	707	773	726	660	794	1,054	1,000	2,742	6,461	1,256
Butte	1,514	794	813	1,088	1,114	1,528	974	1,860	2,933	4,997	3,764	1,929
Cleveland	1,527	1,065	1,013	1,102	1,088	1,198	1,281	1,829	3,024	4,470	3,574	2,177
Dallas	1,894	1,062	1,184	996	879	906	1,110	1,218	2,508	4,297	3,989	1,878
Des Moines	1,088	1,037	994	778	932	697	1,106	1,384	1,909	3,685	3,660	1,714
Erie	989	746	958	925	805	890	1,171	1,561	2,459	3,697	3,306	1,574
Indianapolis	1,400	892	1,360	1,214	1,240	1,328	1,468	1,624	2,561	3,665	3,232	2,298
Lansing	1,435	794	748	790	925	1,016	814	1,595	2,318	3,872	5,650	1,525
Lincoln	1,196	1,012	652	726	658	949	962	1,772	3,202	5,072	4,649	2,106
Little Rock	1,876	734	1,150	945	649	896	1,069	1,864	2,763	4,551	2,841	2,072
Minneapolis	1,373	881	922	827	807	886	957	1,364	2,367	3,581	3,566	1,745
Oklahoma City	2,440	911	991	1,309	1,204	1,238	1,290	1,483	2,623	4,057	4,360	1,949
Peoria	1,357	772	1,008	1,010	845	884	1,018	1,467	2,450	3,737	1,730	1,872
Portland, Me.	2,259	868	833	909	876	867	959	1,568	2,347	4,015	3,666	2,044
Portland, Ore.	1,259	757	840	767	684	870	1,087	1,594	2,444	3,321	3,271	1,494
Providence	1,212	548	909	795	783	885	1,026	1,336	2,580	3,835	3,811	1,912
Racine	985	930	918	1,042	951	898	1,074	2,249	2,663	4,120	5,827	1,539
Richmond	1,959	542	629	579	724	892	1,103	1,176	1,718	3,016	2,104	1,974
Sacramento	1,114	1,381	790	1,187	863	709	751	1,306	2,217	3,453	4,735	1,739
St. Joseph	920	621	501	628	597	784	819	1,578	2,602	2,836	2,527	1,745
St. Paul	944	698	719	958	477	803	715	1,060	2,841	4,802	3,000	1,693
Salt Lake City	1,749	806	885	1,060	912	898	1,246	1,812	2,864	4,314	3,338	1,957
San Diego	2,097	1,072	996	980	736	987	1,273	1,423	2,199	4,065	3,986	1,677
Seattle	1,097	874	940	1,082	1,057	1,212	1,152	1,566	2,581	3,989	4,147	1,734
Springfield, Mo.	1,504	653	662	862	690	725	675	1,145	1,705	3,292	3,633	1,979
Syracuse	1,455	962	776	1,234	1,251	898	1,126	1,374	2,718	3,771	14,278	1,847
Topeka	1,093	704	572	638	755	775	1,283	1,493	2,348	3,914	2,650	1,683
Trenton	991	721	905	802	532	778	684	887	2,110	3,282	5,439	1,349
Wheeling	1,402	680	759	668	693	822	713	1,124	1,557	3,419	4,223	1,306
Wichita	1,818	963	917	864	729	904	1,129	1,499	2,505	3,957	4,801	1,600
Worcester	1,586	836	677	797	684	893	1,190	1,469	1,943	4,081	3,443	2,122

TABLE 35: Coefficients of Variation of the 1929 Incomes of Families for Each 1933 Income Group

Identical Samples: Entire-city

	1933 INCOME GROUP (DOLLARS)											
	No income	*1–249*	*250–499*	*500–749*	*750–999*	*1,000–1,499*	*1,500–1,999*	*2,000–2,999*	*3,000–4,499*	*4,500–7,499*	*7,500 & over*	*All families*
Atlanta	2.031	1.202	.940	1.008	.604	.634	.494	.412	.481	.520	.316	1.092
Birmingham	1.505	1.003	.936	.750	.735	.593	.562	.548	.624	.497	.480	1.057
Boise	1.216	.866	.727	.656	.557	.418	.385	.409	.310	.409	.452	.748
Butte	1.099	.635	.593	.671	.628	.729	.449	.617	.645	.632	.289	.952
Cleveland	1.504	1.062	.839	.776	.692	.625	.537	.573	.609	.521	.287	1.044
Dallas	1.205	1.195	1.045	.767	.580	.509	.490	.422	.599	.585	.332	.891
Des Moines	1.524	1.125	.893	.626	.640	.424	.500	.491	.476	.575	.300	.905
Erie	1.146	.794	.797	.680	.509	.467	.498	.500	.546	.529	.318	.896
Indianapolis	1.586	.953	1.148	.868	.764	.691	.599	.518	.556	.525	.282	1.001
Lansing	1.309	.774	.636	.536	.562	.516	.348	.510	.551	.438	.384	.803
Lincoln	1.353	1.234	.631	.595	.483	.557	.449	.580	.686	.610	.360	1.076
Little Rock	2.260	1.147	1.064	.726	.438	.490	.455	.580	.595	.502	.215	1.101
Minneapolis	1.370	.917	.827	.635	.552	.507	.438	.472	.548	.492	.323	.876
Oklahoma City	1.498	1.099	.845	.892	.714	.654	.551	.512	.590	.546	.437	.948
Peoria	1.325	.919	.958	.763	.595	.520	.488	.504	.559	.491	.142	.964
Portland, Me.	1.844	.871	.772	.734	.634	.540	.455	.535	.534	.527	.303	.979
Portland, Ore.	1.464	.801	.795	.590	.481	.495	.490	.530	.549	.515	.315	.872
Providence	1.434	.609	.845	.651	.564	.528	.477	.474	.581	.486	.349	1.001
Racine	.984	.815	.650	.675	.558	.470	.454	.672	.562	.511	.617	.860
Richmond	2.455	1.006	.937	.618	.608	.552	.521	.421	.420	.430	.186	1.035
Sacramento	1.184	1.401	.692	.848	.556	.415	.348	.470	.554	.491	.477	.829
St. Joseph	1.458	.876	.641	.568	.498	.491	.395	.532	.597	.370	.239	.977
St. Paul	1.078	.815	.703	.778	.359	.513	.351	.395	.698	.701	.185	.960
Salt Lake City	1.670	.895	.792	.773	.600	.506	.554	.602	.644	.584	.255	1.001
San Diego	1.856	1.261	.917	.792	.543	.595	.584	.501	.527	.594	.362	.906
Seattle	1.037	.909	.792	.768	.662	.627	.502	.532	.584	.544	.395	.892
Springfield, Mo.	1.878	.944	.735	.731	.513	.469	.338	.399	.454	.443	.362	.893
Syracuse	1.545	.955	.700	.865	.800	.500	.489	.466	.632	.562	.648	.936
Topeka	1.495	1.023	.692	.561	.524	.459	.577	.523	.554	.547	.203	.953
Trenton	1.223	.811	.911	.690	.428	.485	.324	.334	.511	.468	.593	.854
Wheeling	2.047	.986	.786	.593	.529	.492	.352	.428	.412	.467	.495	.876
Wichita	1.668	1.019	.778	.645	.489	.484	.499	.503	.554	.578	.568	.864
Worcester	1.696	.950	.704	.666	.492	.533	.545	.533	.473	.526	.285	1.034

dispersion, of the 1929 incomes of families that constituted the various 1933 income groups, are supplemented by the corresponding measures for the 1929 incomes of all families.

With respect to absolute heterogeneity, the situation is similar to that previously found: the high 1933 income groups have the most diverse origins; they are followed by the lowest groups, and finally by the middle and moderately low. The standard deviations describe a less uniform course. Secondary maxima appear in the lower middle groups. Moreover, maximum heterogeneity is frequently observed for the next-to-highest income group, not for the highest.

Relative heterogeneity is greatest in the lowest two income groups. It declines as we go up the income scale; then rises, in most cases, to a secondary maximum. In contrast to the pattern of heterogeneity in the 1929 groups, the secondary maximum usually does not occur in the highest group; in many cases it appears in the $3,000-4,499 group.

Apparently the most important difference between the results of this and the foregoing analysis relates to the top income groups. In comparison with the low and moderate income groups, they show a less definite excess of absolute heterogeneity and actually less relative heterogeneity. Although in terms of 1929 groups the top (and lowest) incomes show greatest heterogeneity with respect to income change, in terms of 1933 groups the moderately high (and lowest) groups take the lead. The diversity of the top ranking families of 1933 with respect to their *former* incomes is less conspicuous than the diversity of the top families of 1929 with respect to their *subsequent* incomes.

2 *Groups Showing More or Less Favorable Income Shifts*

a PROBLEMS OF MEASUREMENT

We take up now the second problem posed in the introduction to this chapter: Are all income groups equally affected by the shifts in income rank? Do some experience smaller declines in income—or greater increases—than others? Our interest is focused on the relation between the mean 1929 and 1933 incomes of the families that were in certain income brackets in the base year, either 1929 or 1933. For reasons of economy, this analysis is confined to the entire-city samples.

A standard of comparison must be selected. At first glance, the proportional change in the incomes of all groups might seem a useful standard; that is, we might define a 1929 income group as 'favored' if its 1933 mean income were larger than its 1929 mean income multiplied by the ratio $\frac{\bar{y}}{\bar{x}}$ of the 1933 and 1929 mean incomes of *all* families in 1933 and 1929, respectively. If no groups were 'favored' or 'disfavored', the mean (y_i) and 1929 (x_i) incomes of every 1929 group i would be related as follows:

$$(1)\ y_i - \bar{y} = \frac{\bar{y}}{\bar{x}} (x_i - \bar{x}).$$

This standard of comparison was not adopted, for deviations from it represent a hodge-podge of phenomena, some of which are not at all related to shifts in income rank proper, while others merely indicate that there are shifts of some sort in the distribution as a whole.

Suppose income inequality changes, e.g., the coefficient of variation increases from 1929 to 1933, while income rank does not change at all; i.e., the correlation of 1929 and 1933 family incomes is +1. The relation between the 1929 and 1933 income of the 1929 classes will then be:

$$(2)\ (y_i - \bar{y}) = \frac{\sigma_y}{\sigma_x} (x_i - \bar{x});\ \frac{\sigma_y}{\sigma_x} > \frac{\bar{y}}{\bar{x}}.$$

While the standard case (1) is represented by a straight line through the origin and the center of gravity ($\bar{y}$; $\bar{x}$), situation (2) will find the mean income points (y_i; x_i) along a steeper line through the center of gravity, indicating that the lower income groups ($x_i < \bar{x}$) deviate negatively while the higher ($x_i > \bar{x}$) deviate positively. The comparison of the two model cases shows that with a change in dispersion (case 2), various income groups will show means above ('favored' groups) or below ('disfavored' groups) the line of proportional change (case 1), these positive and negative deviations having no connection with changes in income rank. Much as we are interested in changes in inequality, we do not want them to predetermine the outcome of this analysis, which deals with shifts in income rank alone.

Suppose next that inequality, as measured by the coefficient of variation, remains constant while there are some shifts in income rank leading to imperfect correlation of 1929 and 1933 incomes ($r < 1$). In this case, the mean income points may lie on, or be scattered about, a straight line or any other type of line. A straight line fitted to the points by least squares will have the formula:

$$(3)\ (y_i' - \bar{y}) = r\frac{\sigma_y^*}{\sigma_x}(x_i - \bar{x});\ \frac{\sigma_y^*}{\bar{y}} = \frac{\sigma_x}{\bar{x}}.$$

Since by definition $\frac{\sigma_y^*}{\sigma_x} = \frac{\bar{y}}{\bar{x}}$, and since $r<1$, the slope of the line will be less than $\frac{\bar{y}}{\bar{x}}$.[14] The difference between any 1933 mean income y_i and the standard line (1) will be the algebraic sum of two components: (a) the difference between y_i and y_i', the ordinate of line (3), and (b) the difference between y_i' and the ordinate of line (1), all taken at the same value of x (x_i). Chart 19 illustrates the lines (1), (2), and (3), as well as the situation just described.

Of the two components (a) and (b), the first alone relates to the special fate of group i when income ranks shift. The second reflects the shifts and measures their extent in the distribution as a whole. The above discussion explains why standard (1) was rejected in favor of a more satisfactory standard, namely, the straight line fitted to the mean income points by the method of least squares. This line represents the effect of changes in income level, inequality, and shifts in rank in the distribution as a whole. The formula of this line,

$$(4)\ (y_i'' - \bar{y}) = r\frac{\sigma_y}{\sigma_x}(x_i - \bar{x}),$$

is the same as (3) except that the real standard deviation of 1933 income (σ_y) has taken the place of the theoretical one (σ_y^*), i.e., the assumption of constant coefficients of variation has been dropped. In terms of Chart 19, this line will run between lines (2) and (3).[15]

[14] That is to say, if there are shifts in income rank but no change in income inequality, the 1933 mean income of the 1929 groups must diverge less, on the whole, than their 1929 mean incomes. "The fact that individuals 'wander', that their relative income status shifts from one year to the next, means that the 'extreme' 1929 income classes will be less extreme in [1933] than in 1929". (Friedman and Kuznets, *op. cit.*, Ch. 7, note 13.) As a result, the 1933 means of the 1929 groups tend to regress toward the general 1933 mean as compared with the constellation of the 1929 means of these groups about the general 1929 mean. This phenomenon was greatly clarified by Harold Hotelling (Review of Horace Secrist, 'The Triumph of Mediocrity in Business', *Journal of the American Statistical Association,* Dec. 1933, pp. 463–5; discussion, *ibid.*, June 1934, pp. 196–200); and Milton Friedman (*op. cit.*).

[15] Since in the usual case, illustrated by Chart 19, $r < 1$, that is to say,

$$\sigma_y/\sigma_x > \bar{y}/\bar{x}, \text{ we have}$$

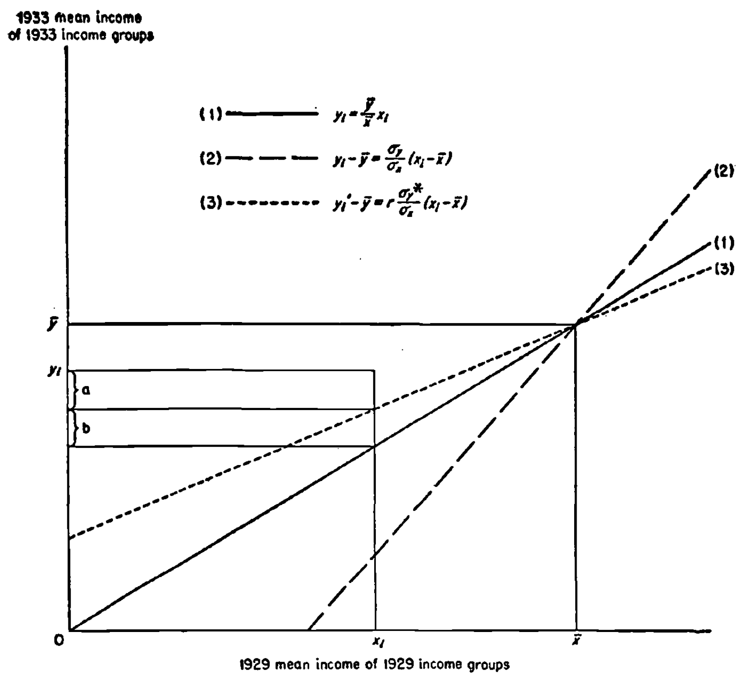
$$r\frac{\sigma_y^*}{\sigma_x} = r\frac{\bar{y}}{\bar{x}} < r\frac{\sigma_y}{\sigma_x} < \frac{\sigma_y}{\sigma_x},$$

and therefore, the slope of line (4) is less than that of line (2) and more than that of line (3).

The deviation of any observed 1933 mean income from line (4) $(y_i - y_i'')$ reflects the particular fate of this group. We may define the measure of that particular fate as the observed 1933 mean income after allowance for the changes in three general characteristics of the entire income distribution: mean level, inequality, and extent of shifts in rank in the distribution as a whole. If $(y_i - y_i'')$

CHART 19

Illustration of Various Regressions



is positive, we may say that group i tends to be 'favored': its 1933 income exceeds the amount that allowance for the three general characteristics would grant. If $(y_i - y_i'')$ is negative, we may say that group i tends to be 'disfavored': its 1933 income falls short of the amount that allowance for the three general characteristics would grant.

Of course, one must hesitate to regard a different $(y_i - y_i'')$ of any size as a valid indication of a positive or negative deviation. The differences between the class means y_i and their standard values derived from (4) may be small enough to be ascribed to random errors. The proper test would entail an analysis of variance for each city sample. For technical reasons it was not attempted here.[16]

Instead, we ask: Is there statistical evidence for a certain pattern of 'favored' and 'disfavored' groups in the various cities? Do the incomes of certain groups tend to deviate positively or negatively in so many cities that a significant pattern is established for the 33 cities as a whole? In other words, we replace the problem of the size of $(y_i - y_i'')$ in particular cities with the problem of the regularity with which positive, or negative, values of this difference appear in a certain income group in all cities. The statistical significance of the regularity of the pattern can be tested by the χ^2 test suggested by W. A. Wallis and G. H. Moore.[17] Thus, when a certain pattern of positive and negative values of $(y_i - y_i'')$ appears with significant regularity for successive income groups in the various cities we call the groups with predominantly positive values 'favored' groups, those with predominantly negative ones, 'disfavored' groups.

In establishing the linear regression of 1933 and 1929 mean incomes for the 1929 groups, the group means were weighted by the reciprocals of their estimated variance.[18] As is obvious from the

[16] The distribution of income differs widely from the normal distribution, not only for the distribution of all incomes in a certain year, but also for the partial distributions of, say, all the 1933 incomes in a 1929 group. For most, particularly the extreme, 1929 or 1933 groups, the distribution of the 'other' year's income is very skew. The analysis of variance is based on the assumption that residuals (e.g., deviations of 1933 family incomes from their mean, within 1929 groups) are normally distributed. While moderate departures from normality do not detract seriously from the efficiency of the test, very pronounced divergences are likely to have that effect.

Facing the analogous situation in their study of professional incomes, Friedman and Kuznets concluded that "little confidence can be placed in the results of tests assuming normality when . . . the data . . . are known to deviate widely from normality" (*op. cit.*, Appendix to Ch. 7, Sec. 1b). Therefore, we refrained from undertaking the laborious task of testing the randomness of the deviations $(y_i - y_i'')$ in the various samples.

[17] The Null-hypothesis is that the signs of the deviations are distributed at random among the various income groups. The city samples are assumed to be drawn from the same population with respect to the effect of the shifts in income rank on the various groups. No assumption of normality is involved. ('Time Series Significance Tests based on Signs of Differences', *Journal of the American Statistical Association,* June 1943.)

[18] The weights (w_i) for the various classes are:

$$w_i = \frac{1}{\sigma_{v_i}^2} = \frac{f_i\,(f_i - 1)}{\sum_{j=1}^{11} f_{ij}\,[(y_{ij} - y_i)^2},$$

where f_i = the number of families in the i-th 1929 class, y_i = their mean income in 1933 estimated on the basis of the horizontal 1933 means, f_{ij} = the number of families in the cell formed by the i-th 1929 class and j-th 1933 class, y_{ij} = their mean income as estipreceding section, the standard deviations of 1933 income within 1929 groups vary considerably and systematically from group to group, while the cross-classification tables in Appendix B reveal the great diversity of the number of families in the various groups. Thus, the reliability of the 1933 means from a sampling point of view is not the same for each group. The complete weighting procedure is required for correct fitting of the least square lines.[19]

In the same fashion we may seek to determine whether certain 1933 income groups are favored or disfavored in an imaginary reverse movement from 1933 to 1929; or, inverting the signs obtained in such an analogous analysis, we may learn whether the various groups of *families that were scattered in 1929 but converged on certain 1933 income groups* are favored or disfavored with respect to the standard of type (4). In the present case, the standard assumes the form:

$$\text{(4a)}\quad (x_j'' - \bar{x}) = r\,\frac{\sigma_x}{\sigma_y}\,(y_j - \bar{y}),$$

where y_j and x_j are the 1933 and 1929 mean incomes respectively of the 1933 income class j, and where the regression is computed on the basis of weighted 1933 class means.[20]

Note 18 Concluded:

mated by the mean income of the entire j-th class in 1933 (y_j). The estimate of the variance of an individual 1933 income in class i is:

$$\sigma_y^2 = \frac{\sum_{j=1}^{11} f_{ij}\,(y_{ij} - y_i)^2}{f_i - 1},$$

that of the group mean of i in 1933: $\sigma_{y_i}^2 = \frac{\sigma_y^2}{f_i}$. See App. A 2.

[19] In computing the correlation coefficients in Table 30, the class means were weighted by the factors f_i only. If the complete weighting system had been used, the coefficients of correlation would have differed according to the base year; i.e., instead of each coefficient there would have been a pair, one involving weights based on 1929 groups, the other, weights based on 1933 groups. This would have meant unnecessary inconvenience. It must be realized, however, that the correlation coefficients (in Table 30) are not quite consistent with the fitted regression lines discussed in this chapter. Completely weighted' correlation coefficients, not shown here, correspond to these lines.

[20] The weights are:

$$w_j = \frac{1}{\sigma_{x_j}^2} = \frac{f_j\,(f_j - 1)}{\sum_{i=1}^{11} f_{ij}\,(x_{ij} - x_j)^2},$$

where f_j = the number of families in the j-th 1933 class, x_j = their mean income in 1929 estimated on the basis of the vertical 1929 means, f_{ij} = the number of families in the cell formed by the i-th 1929 and the j-th 1933 class, x_{ij} their mean income as estimated by the mean income of the entire-city i-th class in 1929 (x_i).

b RESULTS

Straight lines were fitted to the mean incomes of the 1929 and 1933 income classes in 1929 and 1933. Their constants, i.e., their intercepts and slopes, are shown in Table 36. Although, on the whole, these lines fit the observations well, certain patterns of deviation appear.

The analysis of the 1929 group incomes shows that the lowest ('no income') group deviates positively. In all cities the 1933 income of this group is above the standard value, which represents the total effect of changes in income level and dispersion and the frequency and extent of the shifts in income rank on the distribution as a whole. The next-to-lowest group ($1-249) deviates positively in most (24) cities, negatively in a minority (9). The two groups between $250 and $749 do not show a clear excess of positive or negative deviations, but the following two groups, between $750 and $1,499, show negative deviations in 27 and 26 cities, respectively. The $1,500-1,999 group shows no clear preponderance of either positive or negative deviations; but the $2,000-2,999 group deviates positively in all cities, and the next in most (27) cities. The $4,500-7,499 group shows no prevalence of either type of deviation. The group at the top of the 1929 distribution deviates negatively in all cities except one. The signs of the deviations ($y_i - y_i''$) are shown in Table 37.

Thus, there appears a fairly regular pattern in the form of a wave-like movement about the various regression lines. The two very lowest 1929 income groups and the upper central groups, $2,000-4,499, tend to be 'favored', the lower central groups and the top group to be 'disfavored'. From a statistical viewpoint, the regularity of the pattern is beyond doubt. During the Great Depression the recipients of very low 1929 incomes experienced an actual rise in mean income, well in excess of what could be ascribed to the shifts in income rank in the distribution as a whole (the 'regression toward the mean'). Those of moderately high 1929 incomes experienced a relatively small decline in mean income. On the other hand, the mean incomes of both the recipients of moderately low and highest 1929 incomes show rather large declines.

These observations may be explained in terms of the factors held responsible for the changes in income dispersion (see Ch. 2, Sec. 4). The high flexibility of income from property as a whole may have produced the relatively big decline in the top incomes of 1929, while

TABLE 36

Constants of Weighted Regression Equations
Identical Samples: Entire-city

	BASED ON 1929 INCOME GROUPS		BASED ON 1933 INCOME GROUPS	
	Intercept (dollars)	Slope	Intercept (dollars)	Slope
	(1)	(2)	(3)	(4)
Atlanta	89	.641	448	1.055
Birmingham	110	.461	595	1.177
Boise	181	.608	635	.834
Butte	160	.430	1,126	.750
Cleveland	214	.487	809	.983
Dallas	221	.583	722	.925
Des Moines	159	.611	678	.890
Erie	119	.474	808	.948
Indianapolis	153	.595	752	1.018
Lansing	214	.441	866	.930
Lincoln	204	.560	608	.968
Little Rock	94	.569	484	1.174
Minneapolis	202	.580	704	.924
Oklahoma City	229	.518	804	.937
Peoria	227	.551	652	.936
Portland, Me.	200	.618	635	.920
Portland, Ore.	164	.510	734	.898
Providence	202	.579	659	.901
Racine	187	.375	1,006	.837
Richmond	112	.721	274	1.096
Sacramento	219	.616	771	.841
St. Joseph	168	.645	451	.966
St. Paul	96	.655	646	.837
Salt Lake City	194	.540	705	.959
San Diego	295	.548	629	.915
Seattle	259	.481	803	.936
Springfield, Mo.	132	.590	563	.895
Syracuse	164	.554	787	.904
Topeka	140	.550	499	1.086
Trenton	152	.508	631	1.043
Wheeling	189	.530	581	.896
Wichita	181	.516	778	.906
Worcester	225	.590	588	.961

General equation: $y'' = a + bx;\ x'' = a + by$, where x and y = income in 1929 and 1933, respectively, x'' and y'' = regression line values of 1929 and 1933 income, respectively, a = intercept and b = slope.

TABLE 37: 1929 Income Groups

'Favored' (+) and 'Disfavored' (−) Groups, Signs of Deviations from Weighted Regression Line ($y'' = a + bx$)

Identical Samples: Entire-city

	1929 INCOME GROUPS (DOLLARS)											
	No income (1)	*1–249* (2)	*250–499* (3)	*500–749* (4)	*750–999* (5)	*1,000–1,499* (6)	*1,500–1,999* (7)	*2,000–2,999* (8)	*3,000–4,499* (9)	*4,500–7,499* (10)	*7,500 & over* (11)	*No. of plus signs* (12)
Atlanta	+	+	−	−	−	−	+	+	+	−	−	5
Birmingham	+	+	−	−	−	−	+	+	+	−	−	5
Boise	+	+	−	−	−	−	+	+	−	+	−	5
Butte	+	+	+	+	−	−	−	+	+	+	−	7
Cleveland	+	+	+	−	−	−	−	+	+	+	−	6
Dallas	+	−	−	−	−	−	+	+	−	−	−	3
Des Moines	+	+	−	−	−	−	+	+	−	−	−	4
Erie	+	+	+	+	−	−	+	+	+	+	−	8
Indianapolis	+	−	−	−	+	+	−	+	+	−	−	5
Lansing	+	+	+	+	−	−	−	+	+	−	+	7
Lincoln	+	+	+	−	−	−	−	+	+	+	−	6
Little Rock	+	−	−	−	+	+	+	+	+	−	−	6
Minneapolis	+	+	+	−	−	−	+	+	+	+	−	7
Oklahoma City	+	−	−	+	−	+	+	+	+	−	−	6
Peoria	+	−	−	−	−	−	−	+	+	+	−	4
Portland, Me.	+	+	−	+	−	+	+	+	+	−	−	7
Portland, Ore.	+	+	−	+	−	−	+	+	+	−	−	6
Providence	+	−	+	+	+	−	−	+	+	+	−	7
Racine	+	+	−	+	+	−	−	+	+	+	−	7
Richmond	+	+	−	−	−	+	+	+	+	−	−	6
Sacramento	+	+	+	−	−	−	−	+	+	−	−	5
St. Joseph	+	+	−	−	+	−	+	+	−	+	−	6
St. Paul	+	+	+	+	−	−	+	+	+	−	−	7
Salt Lake City	+	+	+	−	−	−	+	+	+	+	−	7
San Diego	+	−	−	−	−	+	+	+	−	−	−	4
Seattle	+	−	−	−	−	−	+	+	+	−	−	4
Springfield, Mo.	+	−	−	+	−	−	−	+	+	−	−	4
Syracuse	+	+	+	−	−	−	+	+	+	−	−	6
Topeka	+	+	−	−	−	−	−	+	−	−	−	3
Trenton	+	+	+	+	−	−	−	+	+	+	−	7
Wheeling	+	+	+	−	−	−	+	+	+	+	−	7
Wichita	+	+	+	−	−	+	+	+	+	−	−	7
Worcester	+	+	+	+	+	−	−	+	+	+	−	8
No. of plus signs	33	24	15	12	6	7	20	33	27	14	1	192 = P
Squares	1,089	576	225	144	36	49	400	1,089	729	196	1	4,534 = S

No. of samples (z) = 33 No. of classes (k) = 11 $\chi^2 = \frac{zk\,(ks - P^2)}{P\,(zk - P)} = 144$ D.o.F. = 10

Probability of reaching or exceeding this value of χ^2 by chance = $<$ .00001.

the direct and indirect effects of unemployment may account for the relatively large drop in the moderately low 1929 incomes, most of which went, in all probability, to wage earners. The mild decline in the moderately high incomes may be ascribed to the relative stability, in terms of the income of salaried and skilled wage earners, which is explained, at least in part, by the milder incidence of unemployment among the recipients of salaries and higher wages. As for the large increase in the very low 1929 incomes, no explanation in terms of the factors previously discussed can be suggested. It may perhaps be due to a very high proportion of young people in that group; entering employment in 1929 as helpers or casual workers, they were able, with growing economic maturity, to raise their incomes relatively to others, despite the depression.

The decline in the mean income of the top 1929 incomes seems to support our explanations in Chapter 2 of the changes in inequality within the upper group. It can be measured by the size of the difference, $y_{11} - y_{11}''$, observed for the eleventh (top) income class ($7,500 and over) in the various cities. As noted, the difference is negative in 32 cities, positive in one. The greater the negative difference in a city, the more, we may say, did the top class lose in relative rank. If the arguments in the preceding paragraph are valid, we should expect that the top incomes would have a declining share in upper-group income when they fell much in relative rank, an increasing share when they fell little or even rose.

The size of the differences $y_{11} - y_{11}''$ was measured, and the cities divided into three groups of eleven each: those where the top group lost most, the negative difference amounting to more than $1,800; those where the top group lost moderately, the negative difference ranging from $800 to $1,800; and those where it lost least, the difference ranging from plus $151 (Lansing) to minus $800 (Table 38).

There appears to be a very pronounced correlation. In all 11 cities where the top group lost most, inequality within the upper group declines. Among the 11 cities of the middle group, declines appear in 9, increases in 2. In the group that lost least or gained, we find only 4 declines, but 7 increases in inequality.[21]

We turn now to the analysis of the incomes of the various 1933 groups. The signs in Table 39 are the *inverted* signs of the deviations $(x_j - x_j'')$. While the original signs indicate whether a group of 1933

[21] X^2 has the value 12; D.o.F., 2. The probability of reaching or exceeding this value by chance is less than 1 per cent.

incomes deviates positively or negatively in an imaginary reverse movement from 1933 to 1929, the inverted signs are consistent with the true direction of the time flow; they indicate whether the group of families that in 1929 were scattered over various groups and then proceeded to form a common 1933 income group were 'favored' or 'disfavored' with respect to other groups of this type.

In almost all cities the group that later was to form the 1933 'no income' group deviates negatively. In less degree the same is true of those converging on the three next-to-lowest groups ($1-749). The three lower central groups ($750-1,999) of 1933 are composed of

TABLE 38

Degree to which Top 1929 Group is 'Disfavored' and Change in Inequality within the Upper Income Group, 1929–1933

	CITIES IN WHICH TOP 1929 GROUP			
	Lost most	Lost moderately	Lost least or gained	All samples
Coefficient of concentration for *u* group				
Declines	11	9	4	24
Increases	0	2	7	9
All samples	11	11	11	33

The top 1929 group consists of families with incomes of $7,500 or more. The degree to which it is lost is measured by the size of the deviation of the 1933 mean income of the group (y_{11}) from the value of the y_{11}''. For the range of the three groups of cities, see text.

groups that deviate positively, the three upper central groups ($2,000-7,499), of groups that deviate negatively. For the highest groups the cases are almost equally divided between positive (16) and negative (17) deviates.

Again a fairly regular pattern that is statistically acceptable emerges; but its shape differs from that of the 1929 income groups. The families that come to populate the lowest 1933 income groups are, very probably, the victims of unemployment. Their incomes decline more rapidly than those of all families taken together, even after changes in income dispersion and the shifts in rank are allowed for. Many of the families converging on the moderately high 1933 income groups are probably victims of the collapse of dividends and stock values, families that formerly derived high incomes from these sources. Finally, those who come to form the moderately low 1933

TABLE 39: 1933 Income Groups

'Favored' (+) and 'Disfavored' (−) Groups, Inverted Signs of Deviations from Weighted Regression Line ($x'' = a + by$)

Identical Samples: Entire-city

	1933 INCOME GROUPS (DOLLARS)											
	No income (1)	*1–249* (2)	*250–499* (3)	*500–749* (4)	*750–999* (5)	*1,000–1,499* (6)	*1,500–1,999* (7)	*2,000–2,999* (8)	*3,000–4,499* (9)	*4,500–7,499* (10)	*7,500 & over* (11)	*No. of plus signs* (12)
Atlanta	−	+	+	+	+	−	−	−	−	−	+	5
Birmingham	−	+	−	−	−	+	+	+	−	−	+	5
Boise	−	+	−	−	+	+	−	−	+	−	−	4
Butte	−	−	+	−	+	−	+	−	−	−	−	3
Cleveland	−	−	−	+	+	+	+	−	−	−	+	5
Dallas	−	−	−	+	+	+	+	−	−	−	+	5
Des Moines	−	−	−	−	−	+	−	−	−	−	+	2
Erie	−	+	−	−	+	+	+	−	−	−	+	5
Indianapolis	−	−	−	+	+	+	+	+	−	−	+	6
Lansing	−	−	+	−	+	−	+	−	−	−	+	4
Lincoln	−	−	−	−	+	+	+	−	−	−	−	3
Little Rock	−	+	−	−	+	+	+	−	−	−	−	4
Minneapolis	−	−	−	−	+	+	+	−	−	−	−	3
Oklahoma City	−	+	−	−	−	+	+	+	−	−	−	4
Peoria	−	−	−	−	+	+	+	−	−	−	−	3
Portland, Me.	−	−	−	−	+	+	+	−	−	−	−	3
Portland, Ore.	−	−	+	−	+	+	+	−	−	−	+	5
Providence	−	−	−	−	+	+	+	−	−	−	−	3
Racine	+	+	−	−	+	+	+	−	−	−	−	5
Richmond	−	−	+	−	+	−	+	+	−	−	+	5
Sacramento	−	−	−	−	−	+	+	−	−	−	+	3
St. Joseph	−	−	+	−	+	+	+	−	−	−	+	5
St. Paul	−	−	−	+	+	+	+	−	−	−	−	4
Salt Lake City	−	−	−	−	+	+	+	−	−	−	−	3
San Diego	−	−	−	−	+	+	+	−	−	−	−	3
Seattle	−	+	−	−	+	+	+	+	−	−	−	5
Springfield, Mo.	−	+	−	−	−	+	+	−	−	−	+	4
Syracuse	−	−	+	−	+	+	+	−	−	−	−	4
Topeka	−	−	+	−	−	+	−	+	−	−	+	4
Trenton	−	−	−	−	+	+	−	−	−	−	−	2
Wheeling	−	+	−	−	+	−	+	+	−	−	+	5
Wichita	−	−	−	−	+	+	+	−	−	−	+	4
Worcester	−	−	−	−	+	+	+	+	−	−	−	4
No. of plus signs	1	10	8	5	27	28	28	8	1	0	16	132 = P
Squares	1	100	64	25	729	784	784	64	1	0	256	2,808 = S

No. of samples (z) = 33 No. of classes (k) = 11 $\chi^2 = \frac{zk(ks - P^2)}{P(zk - P)} = 160$ D.o.F. = 10

Probability of reaching or exceeding this value of χ^2 by chance = $<$ .00001.

income groups probably include many wage earners who remained employed during the depression and enjoyed relatively stable wages. Although the identities of the three groups cannot be established beyond doubt with the material at hand, the preceding explanations of income changes seem to be confirmed.