

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: Appendix A: Information about the Data

Chapter Author: Horst Mendershausen

Chapter URL: <http://www.nber.org/chapters/c5311>

Chapter pages in book: (p. 121 - 140)

APPENDIX A

Information about the Data

1 *Nature of Data Used in This Study*

NAME AND AGENCY IN CHARGE OF STUDY AND TABULATIONS	GEO- GRAPHIC COVERAGE	YEARS COVERED	INCOME RECEIVING UNIT	INCOME CONCEPT	SAMPLE SIZE	SOURCE
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Financial Survey of Urban Housing: Dept. of Commerce	33 cities	1929, 1933	Families & house- keeping individuals	Total family cash income	5-13% of occupied dwelling units in city (Table 1, col. 10)	Dept. of Commerce: <i>Finan- cial Survey of Urban Hous- ing</i> ; Wickens: <i>Residential Real Estate</i> & special tabulations.
Consumer Purchases Study: Dept. of Labor	Atlanta, Ga. & Provi- dence, R. I.	12 months ending Dec. 31, 1935-	Families ('compre- hensive sample')	Total family income	About 4% and 20% respectively of addresses listed in city directories of Atlanta & Providence	Dept. of Labor, Bureau of Labor Statistics, Consumer Purchases Study, Bulletins 647, 1, & 645, 1; Washington, 1939
Delaware State Income Tax: U. of Delaware	State of Delaware	Dec. 31, 1936 1936, 1937, 1938	Individuals or married couples	Total income excl. capital gains	Tax returns from almost all income recipients in the state, citizens & resi- dents. 1936: data from all tax returns. Estimates for the following two years based on samples. 1937: 50% sample of taxable returns with 1938 net incomes above \$3,000, 5% samples of other tax- able & nontaxable returns; 1938: 100% sample of taxable returns with 1938 net incomes above \$3,000, 10% samples of other returns	Communication from Lawrence Brokate
Wisconsin State Income Tax: Wisconsin Tax Commission	State of Wisconsin	1929, 1934, 1935, 1936	Individuals or married couples with net taxable incomes of \$800 (\$1,600) or more, & receivers of smaller incomes if requested to report	Net taxable income	1929, '35, '36: complete coverage of persons subject to State Income Tax Law (with their dependents, 30-40% of the state population). 1934: com- plete coverage of total incomes of \$3,000 & more plus inflated 10% sam- ple of returns with lower incomes	Wisconsin Tax Commission: <i>Wisconsin Individual Income Tax Statistics</i> , 1929, 1934, 1935, 1936, Vol. I of each year. Table I
Old-Age & Survivors Insurance: Social Security Board	United States	1937, 1938, 1939, 1940	Workers in 'covered industries'	Taxable wages in 'covered employ- ment'	About 32,000,000 workers receiving in 1938, 65% of aggregate net wages & salaries in the United States. 100% samples for 1937 & 1938, 20% for 1939, 4% for 1940	<i>Handbook of Old-Age and Survivors Insurance Statistics</i> , Washington, 1940, Tables 3 & 8 & communications from Social Security Board
German Income and Wage Tax: Statistisches Reichsamt	Germany (1936: incl. Saar territory)	1936, 1938, 1939, 1934, 1936	Individual income receivers	Total income omit- ting all except short- term (3 months) capital gains & minor items (see <i>Wirtschaft und Statistik</i> , Sept. 1939, p. 560)	Apparently complete coverage (about 31 million income receivers)	<i>Statistisches Jahrbuch für das Deutsche Reich</i> , 1938, p. 5 ^a

2 *Preparation of the Financial Survey Tables for Analysis*

The original tables distinguish 39 income groups for 1929¹ and 11 income groups for 1933. By combining several of the 1929 groups and setting up an identical set of groups for the two years, computation was economized and more compact correlation tables made. Moreover, since information obtained in 1934 on 1929 income is likely to be less accurate than information on income in the year just preceding (Sec. 3), there is not much sense in presenting the 1929 distributions in more detailed form. The 11 income groups used for both 1929 and 1933 are:

No Income ²	\$1,000-1,499
\$ 1-249	1,500-1,999
250-499	2,000-2,999
500-749	3,000-4,499
750-999	4,500-7,499
	7,500 & over

When the income groups are combined, the jaggedness of the distribution disappears (Chart A 1).

Tables similar to the cross-classification tables for the entire-city samples (App. B) were made for the identical samples of all tenants and all owners in each city. For the subsamples by type of canvass, only the marginal totals were computed.

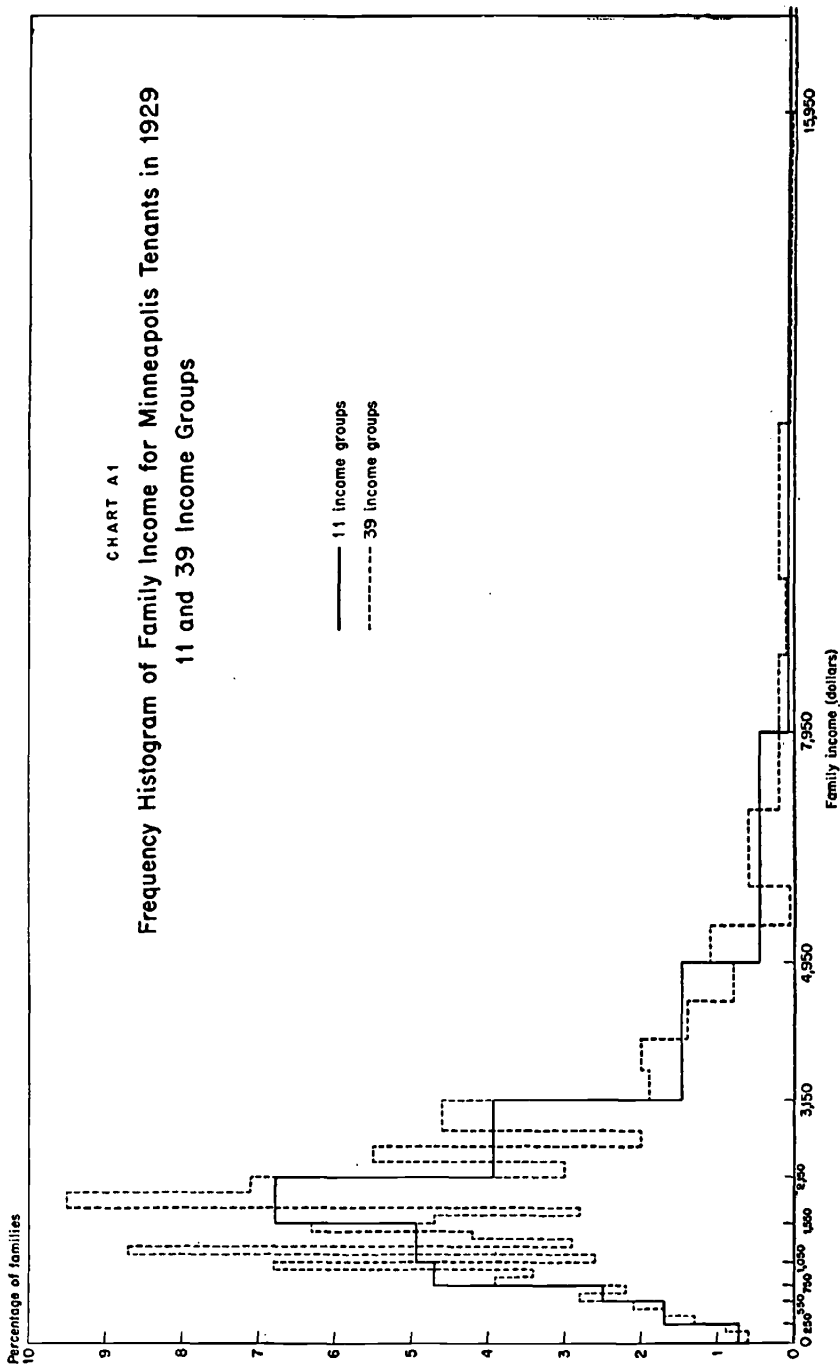
The 'all tenants' and 'all owners' samples were obtained by lumping together the subsamples for the two types of canvass. Among the tenant samples, returns from personal enumeration are more numerous than mail returns in 20 cities, among owners, in 14 cities; but there appears no extreme preponderance of either type of return. In 24 tenant (27 owner) samples each type is between 33 and 66 per cent of all returns, and in 32 tenant (32 owner) samples, between 30 and 76 per cent. In Worcester alone, where the Financial Survey was not completed as planned, is the percentage of mail returns as high as 86 for tenants, 83 for owners.

In comparison with the mail return samples, the personal enumeration samples show as a rule larger relative frequencies in the low income brackets (below \$1,000) and smaller relative frequencies in the high income brackets (above \$4,500). Well-to-do families were somewhat reluctant to state their incomes to a local enumerator, while poor fami-

¹ No income, \$1-149, 16 \$100 intervals up to \$1,749, 5 \$200 intervals up to \$2,749, 3 \$400 intervals up to \$3,949, 4 \$500 intervals up to \$4,949, 4 \$1,000 intervals, and 3 \$2,000 intervals up to \$15,949, \$15,950-19,949, \$19,950 and over.

² The 'no income' group consists almost entirely of unemployed. In 21 cities, for which data are available in the *Financial Survey of Urban Housing*, tenant (owner) families in this group reported that the principal earner was employed; in 1933, less than 6(15) per cent of his normal full working time.

CHART A1
 Frequency Histogram of Family Income for Minneapolis Tenants in 1929
 11 and 39 Income Groups



lies responded more readily to the personal interview which involved neither reading instructions nor writing.

In the absence of any rational basis for the construction of a weighting system, the two types of sample were simply added. Since the two types of sample distributions, regardless of differences in their characteristics for the same year, show very much the same patterns of income change, 1929-33, erroneous weighting is not likely to influence the outcome of the analysis of the income change.

In all cities, except Providence, the identical sample for tenants contains a smaller percentage of the tenant population than the owner sample does of the owner population (Table 1, col. 13). In Providence the percentages are the same. The preponderance of owners, appearing first in the aggregate samples of most cities, tends to increase as schedules are rejected and identical samples formed: as a rule, owners gave more complete information than tenants. Since the proportions of the two tenure groups in the 1933 city populations are known, overrepresentation of owners in the entire-city samples can easily be avoided: the identical tenant samples were inflated to make their sampling ratio equal that for owners. After that correction, the two tenure samples were lumped together. Consequently the 'entire-city' samples contain tenants and owners in the correct (population) proportions. The correction factors for tenants (Table A 1) are derived from Table 1, column 13. We have no way of knowing and of correcting for variations in sampling ratios with respect to other important income determinants, e.g., nativity and occupation.

It is possible to check the proportion of white families in the Financial

TABLE A 1

Correction Factors Applied to Tenant Samples

Atlanta	1.18	Providence	1.00
Birmingham	1.56	Racine	1.68
Boise	1.55	Richmond	1.30
Butte	1.77	Sacramento	1.89
Cleveland	1.43	St. Joseph	1.72
Dallas	1.95	St. Paul	1.23
Des Moines	1.55	Salt Lake City	1.38
Erie	1.56	San Diego	1.77
Indianapolis	1.85	Seattle	1.54
Lansing	1.55	Springfield, Mo.	1.71
Lincoln	1.72	Syracuse	1.42
Little Rock	1.48	Topeka	1.45
Minneapolis	1.27	Trenton	1.34
Oklahoma City	1.54	Wheeling	1.51
Peoria	1.47	Wichita	1.91
Portland, Me.	1.31	Worcester	1.18
Portland, Ore.	1.67		

Survey usable sample (1933) for some cities, e.g., Atlanta and Cleveland.³ For Atlanta, we made a comparison with the record card sample of the Study of Consumer Purchases (1935-36); for Cleveland, with the 1930 *Census of Population*. (The Financial Survey data were corrected for the overweighting of owners in the usable samples.) The Financial Survey covers the metropolitan district, the other data cover the city proper. The comparison reveals a fair degree of correspondence in the ratios.

Percentage of white families

	ATLANTA	CLEVELAND
Financial Survey	63.1	93.1
Consumer Purchases Study	67.1	
Census		91.9

For tenants alone, similar checks were carried out using the Real Property Inventory as the standard of comparison. Again the Financial Survey samples seem rather satisfactory.

Percentage of white families (among tenants)

	ATLANTA	CLEVELAND	SAN DIEGO	SEATTLE
Financial Survey	57.0	91.9	95.6	96.1
Real Property Inventory	63.1	90.6	94.0	97.3

Four sets of mean incomes must be distinguished: (a) the 1929 means of the 1929 income groups, which we call the vertical 1929 means (see arrangement in Appendix B tables); (b) the 1933 means of the 1933 income groups (horizontal 1933 means); (c) the 1933 means of the 1929 income groups (vertical 1933 means), and (d) the 1929 means of the 1933 income groups (horizontal 1929 means). The group means under (a) and (b) were established for all entire-city samples, tenants and owners, mail return and personal enumeration. The group means under (c) and (d) were not established for the subsamples by type of canvass, since these samples were not subjected to the laborious analysis of shifts in income rank.

a) The original tables show the vertical 1929 means for the usable sample of each tenure group and type of canvass in a city. In the absence of specific information it was assumed that these means can be applied to the identical sample. Incomes from both types of canvass were added for each tenure group, group by group, and from the total incomes for each group, the mean incomes were computed. The mean incomes for the entire city were calculated similarly, the difference being that the tenant and owner totals were weighted according to the scheme described above.

b) The horizontal 1933 means, not shown in the original tables, were obtained from the *Financial Survey of Urban Housing* and special com-

³ Sources: *Financial Survey*, pp. 264 ff., 733 ff., 1119, 1182; Department of Labor, Bureau of Labor Statistics, *Family Income in the Southeastern Region, 1935-36*, Bul. 647 (Washington, D. C., 1939), I, 498; *Census of Population*, VI, 56.

munications from the Department of Commerce, both referring to the usable 1933 samples of all tenants and all owners, separately. It was assumed that the means are the same in the two subsamples of a tenure group and that they hold for the identical samples as well. The entire-city means were computed in a way analogous to that followed in the case of the vertical 1929 means.

c) The vertical 1933 means, given in the original tables for usable samples, were treated in the same way as the vertical 1929 incomes and used in the computation of the correlation coefficients between 1929 and 1933 family incomes in tenant, owner, and entire-city samples. However, in the computation of standard deviations and coefficients of variation of 1933 incomes within 1929 groups, the vertical 1933 means of the identical samples were estimated with the help of the horizontal 1933 means for those samples. The total 1933 income of the 1929 income group i (Y_i) is the sum of the products of all cell frequencies in group i (f_{ij}) and the corresponding 1933 mean incomes of the 1933 income groups (\bar{y}_j). The formula is:

$Y_i = \sum_{j=1}^{11} (f_{ij} \bar{y}_j)$, and the mean 1933 income of the 1929 income group i is: $\bar{y}_i = Y_i/f_i$, where $f_i = \sum_{j=1}^{11} f_{ij}$ stands for the total number of families in group i . These estimates furnish consistent sets of horizontal and vertical 1933 means for the computation of the group-wise dispersion measures. This method was used for each of the 33 entire-city samples.

d) Horizontal 1929 means were computed as follows: The total 1929 income of the 1933 income group j (X_j) is the sum of the products of all cell frequencies in group j (f_{ij}) and the corresponding 1929 mean incomes of the 1929 income groups (\bar{x}_i). The formula is:

$X_j = \sum_{i=1}^{11} (f_{ij} \bar{x}_i)$, and the mean 1929 income of the 1933 income group j is: X_j/f_j , where $f_j = \sum_{i=1}^{11} f_{ij}$ stands for the total number of families in group j . These means were established for tenants, owners, and entire cities, separately.

Finally, the 1929 mean income of the families in a certain cell, ij , was estimated to be equal to the vertical 1929 mean of the i -th group, the 1933 mean income of the cell, to the horizontal 1933 mean of the j -th group.

3 Shortcomings of the Financial Survey Income Data

Ideal data for our investigation would satisfy two requirements: (a) They would furnish representative samples of the income receiving units in a city, for both 1929 and 1933. (b) Each individual income figure would represent the correct income of a well defined type earned by an equally well defined and constant income receiving unit. No data available at present on the distribution of incomes by size fully satisfy these require-

ments. In some respects the Financial Survey data are of less satisfactory quality than other materials. For instance, in the Study of Consumer Purchases the concepts were more clearly defined and the sampling procedures more satisfactory. In other respects the Financial Survey furnishes more valuable information: it covers all economic and racial groups and family types, and above all, it shows incomes of the same families in different years.

In connection with requirement (b) it was noted that the Financial Survey did not define 'income' or 'family' precisely. The nature of any bias arising from the application of different income and family concepts in the individual returns cannot be determined; but as the questions were indefinite and the information was not checked, it is likely that many families misstated their income in terms of any single income concept. Probably many underestimates of family income were unwitting, the reporting family member having forgotten or never known certain income items. This source of error is particularly important for 1929 income, since the chance that an item will be forgotten increases with the length of the time between receipt of the money and the interview. In addition, some misstatements may have been deliberate. Families that were poor in 1933 may have believed that understating their incomes would help them to stay or to get on relief rolls. Well-to-do families may have understated their incomes because they were afraid the truth might increase their income tax payments. Both tendencies may have influenced 1933 income statements; neither can have greatly affected 1929 income statements.

These tendencies to understate income may have been offset to some extent by overstatements of 1933 and perhaps even 1929 incomes in order to 'show off', and of 1929 incomes because of too rosy memories of prosperity days.

The total earned income of a family during a given year is determined by the number of gainful workers in the family, their rate of remuneration per unit of time, and the time worked. The second and third factors were subject to considerable variation between 1929 and 1933. Our interest centers upon the composite effect of their variation. It is deplorable that we cannot be sure of the constancy of the first factor. The incomes of individuals who were members of the family proper in 1933 but not in 1929 or conversely may have been included in family income in one or both years. Thus the number of gainful workers per family may have varied, and consequently family income. Or, the income of a family that doubled-up with another family during the depression may have been included in the 1933 income of the latter, but not in the 1929 income. Variations of this nature should be excluded as far as possible in retrospective income studies.

To what extent does the present material fall short of requirement (a)?

Sampling the population of a city by entire blocks is probably cheaper than drawing addresses at random from a city directory, because interviewers spend less time moving from one dwelling unit to the next, but it produces a larger random sampling error. The sample units, i.e., the families, are not selected regardless of income; only some superunit, the block, is chosen on that basis. Within individual blocks family income varies less than in the city as a whole, since people with similar incomes tend to live in the same neighborhood.

Under the extreme assumption that all families within a block have the same income, the number of independent observations in a sample of 10,000 families drawn from 100 random blocks would not be 10,000 but 100 and the sample standard error would be ten times as large as in the case of sampling by individual families. Therefore the true number of independent observations in our samples must be smaller than the number of families covered; how much, we do not know.

While sampling by blocks has tended to increase the importance of random error in the sample distributions, the instruction given city supervisors to add blocks whenever they felt that the sample was 'not representative' may have introduced some bias in the distributions.

With these and the previously noted shortcomings in mind we are not astonished to find the samples somewhat lacking in representativeness. While there is no material for checking the samples against the distribution of incomes in the universe, it is possible to compare the distributions of tenant families by the rent paid in 1933 (usable sample) with the corresponding distributions in the Real Property Inventory, and similarly for the distributions of owner families by the value of their house on January 1, 1934. Considering the Financial Survey distributions as samples from the universe represented by the Real Property Inventory we find, by χ^2 tests, that, with few exceptions, deviations of the observed size can be expected from random samples with less than 1 per cent probability only.⁴ Six tenant samples (Erie, Lansing, Lincoln, Little Rock, Salt Lake City, and Topeka) are exceptions. Among them the probability that the observed deviations would appear in random sampling is .03, .01, 0.4, .01, 0.2, and .04, respectively. Only two in 66 samples (tenants in Lincoln and Salt Lake City) may be considered unbiased; the rest show systematic deviations from the population rent and value distributions for 1933, that is, deviations that can be expected from random samples with less than 4 per cent probability.⁵

⁴ We used the 7 rent and 9 value classes in the *Financial Survey of Urban Housing* on the first page for each city.

⁵ However, these tests are not adequate; for the readings from the χ^2 tables are based on the assumption that the individual items that enter into the computation of the class frequencies were drawn at random. In the case of the Financial Survey, blocks, i.e., relatively *homogeneous* groups of individual items, sampled at random.

The discrepancies between the two bodies of data show a certain pattern. The Financial Survey frequencies generally appear too small in the lowest and highest rent or value groups, too large in the central groups.⁶ Certain obstacles seem to have prevented the collection of sufficient returns from extreme groups: possibly, ineffective sampling of the poor, reluctance among the well-to-do to give information on their economic status, particularly in a personal interview. Among the rejected (incomplete) schedules the proportion of the extreme groups is large.⁷

Since both apartment rentals and values of owner-occupied residential property are positively correlated with income⁸ we may expect the sample *income* distributions for 1933 to show somewhat excessive relative frequencies in the center, too small frequencies at both ends.

a BIAS IN THE MEAN INCOME

How do the shortcomings of the Financial Survey data affect our comparisons of 1929 and 1933 incomes? There are indications that the samples of 1933 incomes are biased with respect to the 1933 income universe. Since information on 1929 and 1933 incomes was collected at the same time and from the same people some bias may have been imparted to the 1929 data. Even if the samples of 1933 incomes had been unbiased there would be no assurance that the samples of 1929 incomes are unbiased with respect to the 1929 income universe. The incomes of certain groups of families that existed in 1929 were barred from the sample by the very fact that it was taken later and separately for individual cities: after 1929 some families broke up because of death, divorce, or separation; others moved out of the city. None of these families had a chance of being included in the samples. Consequently, the present samples of 1929 incomes may be biased with respect to the 1929 universe. What can be

⁶ Around \$4,000 of value and \$300 of annual rent respectively; see below, Sec. a.

⁷ For 33 cities covered by the Financial Survey the frequencies in the rejected sample with 1933 income data form the following percentages of the 'enlarged income sample' for 1933:

incomes below \$500	12.3
incomes \$500-7,499	8.3
incomes above \$7,500	11.5
all incomes	9.4

The 'enlarged income sample' for 1933 consists of identical sample + half-reporting sample (1933) + rejected sample with 1933 income data.

⁸ For instance, among tenants in Minneapolis the correlation ratio between (1) 1933 income and (2) 1933 rent is $r_{12} = .65$ (11 classes both ways, 9,223 observations). With rising rent, the expected income tends to increase along a path of steadily growing steepness. From an inspection of mean residential values for various income groups it appears that the correlation of value and income can usually be approximated by a straight line. Since values of owner-occupied dwellings and incomes have not been cross tabulated in further detail no correlation was computed.

said about the probable nature of this bias, and what about the resulting bias in differences between 1929 and 1933 incomes?

i) Most of the families that broke up after 1929 because of death must have been old families in 1929. While the incomes of old people are usually less than those of people in the best working age, at least as far as earnings are concerned, the mean income of old people from work may yet lie slightly above the general mean. This is indicated indeed by Old-Age and Survivors Insurance data.⁹

They, however, do not reflect income from other sources than work, which are conceivably higher among older people; nor do they take account of the incidence of unemployment on the various age groups. The Unemployment Census of 1937 gives evidence of a slightly higher

*Totally Unemployed and Emergency Workers,
1937, as a Percentage of
Gainful Workers, Census of 1930*

AGE GROUP	% UNEMPLOYED
15-19 years	26
20-29 "	16
30-39 "	13
40-49 "	15
50-59 "	17
60-74 "	17
<hr/>	
15-74 . "	16

Census of Partial Unemployment, Unemployment and Occupations (Washington, D. C., 1938), I, 2. *Census of Population, 1930*, V, 116.

proportion of unemployed in the older age bracket. In a year of large unemployment the mean income of old wage and salaried workers, employed and unemployed, might well be less than the mean income of all workers. This conclusion would be consistent with data for Michigan, 1934, reproduced by Maurice Leven.¹⁰

⁹ Wages of workers in industries covered by the old-age and survivors insurance provisions of the Social Security Act. *Handbook of Old-Age and Survivors Insurance* (Washington, D. C., 1940), pp. 2-5.

*Mean Taxable Wages of Workers Covered by
Old-Age and Survivors Insurance, 1938*

AGE GROUP	MEAN TAXABLE WAGES
Under 19 years	\$209
20-29 years	636
30-39 "	1,024
40-49 "	1,140
50-59 "	1,094
60 years and over	951
<hr/>	
All age groups	846

¹⁰ *The Income Structure of the United States* (Brookings Institution, 1938), p. 156.

But for a prosperous year like 1929 the omission of old families from the distribution probably caused a downward bias in the mean income of the identical sample.

ii) Families that had been set up after 1929 were excluded from the identical sample. Most of them must have been young families, with 1933 incomes probably well below the general mean. Both the wage and unemployment statistics shown above indicate that the income status of young people is below average: they are particularly subject to unemployment, and when employed, receive relatively low wages. Their omission probably imparted an upward bias to the 1933 mean income of the identical samples.

iii) Certain families that lived in a particular Survey city in 1929 moved before the Financial Survey was taken. If their incomes did not closely approximate the general distribution, the resulting bias would not be compensated for in the 1933 sample by the omission of families with similar characteristics moving *into* the Survey city after 1929, for the latter were included in the sample. To the extent that the 'emigrating' families became 'immigrants' in another Survey city, the bias might cancel over some aggregate of Survey cities; but it would remain present in each individual city. Still, we could neglect this source of bias were it possible to assume that, for each individual city, immigration and emigration offset each other with respect to the number of families involved and the course of their income. Such an assumption seems, however, hardly acceptable. There are reasons to believe that during the depression the migratory movement affecting large cities was mainly outward bound, i.e., that there were more 'emigrants' than 'immigrants'.¹¹

Several considerations would suggest that the 1929 incomes of the 'emigrating' families were below the population mean of that year. The migrant group is likely to have comprised a larger proportion of tenants than of owners, or of wage earners than of salaried persons; for it is easier, economically and psychologically, to give up a flat than one's own house, and the compulsion to move is likely to be strongest on the groups most exposed to unemployment. Moreover, the younger age groups form a large proportion of the migrant population. Thus the social, occupational, and age composition of the prospective migrants of 1929 probably caused it to be heavily weighted with low-income families. However, migration presupposes the possession of some funds. Many of the families that entered the depression years with low incomes may have lacked

¹¹ This does not seem to be true for all cities, at least not in 1933, when conditions began to improve. The family population of the Cleveland metropolitan district, for instance, showed a net increase as a consequence of migration: 1.1 per cent of the families left the district and a number corresponding to 1.7 per cent entered. See H. Whipple-Green, *Real Property Inventory of Metropolitan Cleveland*, Report 7 (Cleveland, 1936), p. 13.

such funds. Thus poverty in 1929 may have kept families from moving during the subsequent depression.

Unfortunately there is almost no statistical information on the income distribution of migrating families. Some data for Cleveland indicate that in 1933, at the end of the depression, the well-to-do were more mobile than the poor, both within and beyond the limits of the metropolitan district.¹² Although it is by no means certain that the well-to-do migrants of 1933 were well-to-do in 1929, this bit of information does not favor the hypothesis that the 1929 mean incomes of the prospective migrants was below the population mean. In conclusion, it is hard to say what sort of a bias the omission of the 'emigrating' families imparted to the mean 1929 incomes of the samples.

iv) Some families disintegrated after 1929 as a consequence of divorce or separation. To the extent that the divorced formed separate 'families' they may have been covered by the Financial Survey; but if they joined other families (e.g., parents), their income may have been included in the 1933 income of those families but not in the 1929 income, thus leading to an underestimate of 1929 income as compared with 1933 income.

v) A similar effect may be expected from the doubling-up of several families that often took place during the depression.¹³

To sum up: while it is impossible to assess definitely the net bias that these and other factors have imparted to the sample income distributions, the preceding argument suggests (1) a downward bias in the income level of the 1929 sample with respect to the 1929 population (points a-i, iv, v, in part possibly counterbalanced by point iii) and (2) some upward bias in the income level of the 1933 identical sample with respect to the 1933 population (point a-ii).¹⁴ Both parts of this thesis are supported by a comparison of the Financial Survey tenant families (usable samples) distributed by their 1929 rental bill with corresponding distributions in the *1930 Census of Population*.¹⁵ In the Financial Survey samples we find excessive frequencies in the lower rent groups and deficient frequencies in the higher rent groups. Therefore, it may be expected that the 1929 tenant samples show excessive frequencies in the low income groups.

The table illustrates, for the usable tenant samples, the point made above regarding the deficiency of the 1933 sample frequencies in the

¹² *Ibid.*, p. 76.

¹³ In Cleveland again, October 1932, 12,642 families, or 4.2 per cent of the total, lived together with other families. H. Whipple-Green, *Real Property Inventory of the Cleveland Metropolitan District*, Report 2 (Cleveland, 1933).

¹⁴ This point does not affect the usable samples, which include the new families.

¹⁵ VI, 61 and *passim*. Since the Financial Survey does not furnish a classification of owner-occupied dwelling units according to their value on January 1, 1930, analogous comparisons cannot be made for owners.

*Number of Tenant Samples (out of 32)
Showing Excessive Relative Frequencies
in the Various Rent Classes*

RENT CLASSES	NUMBER OF CITIES IN WHICH THE FINANCIAL SURVEY FREQUENCIES EXCEED THE FREQUENCIES OF THE Real Property	
	1930 Census 1929 sample	Inventory 1933 sample
Under \$120	21	5
120-179	23	13
180-239	26	18
240-359	24	27
360-599	8	19
600-899	4	13
900 & over	2	4

No adequate Census data are available for Boise. The number of tenant samples with smaller relative frequencies is the difference between 32 and the number shown in the table. However, for 1933 some samples show the same frequencies as the Real Property Inventory, i.e., one in the fourth rent class, two in the next to last, and four in the last.

extreme classes.¹⁶ Since both the high and low rent classes find a weaker representation in the Financial Survey than in the Real Property Inventory there is no definite evidence of a bias in the 1933 mean incomes of the *usable* samples. But a tendency toward an upward bias appears in the 1933 mean incomes of the *identical* samples.¹⁷ In 20 (22) cities the identical samples of tenants (owners) show slightly higher mean incomes than the usable samples.

In view of a general downward bias in the mean incomes for 1929 and the prevalence of an upward bias in the 1933 means, we must expect our identical samples to underrate the decline in the income level from 1929 to 1933.

b BIAS IN THE DEGREE OF INCOME INEQUALITY

The degree of income inequality as measured by the coefficients of concentration and of variation is affected by the excessive—or deficient—representation of certain income groups. Therefore, our samples may provide biased estimates of the degree of income inequality in the universe from which they are drawn.

General reasoning suggests that our 1933 samples lead to underesti-

¹⁶ The situation is very similar for owners. Excessive sample frequencies are found in the central value brackets, deficient ones at both extremes.

¹⁷ It may be recalled that for 1929 the mean incomes of all families in the corresponding usable and identical samples coincide in most cases and show very small differences in the others. The most noticeable differences are found for owners in St. Joseph and Trenton, amounting to \$85 and \$82, respectively, in favor of the usable samples.

mates of inequality: underrepresentation of the extreme income groups tends to lessen differences. While there are no data to verify the hypothesis directly, an indirect test may be made by comparing the degree of inequality in the sample and total *rent* distributions of tenants. Our hypothesis is confirmed by the fact that the coefficients of concentration pertaining to the 1933 Real Property Inventory exceed the coefficients of the corresponding Financial Survey distributions for all cities except Birmingham (see Table A 2). It thus appears that our sample *income* distributions for 1933 show too small a degree of inequality.¹⁸

The situation is less clear for the 1929 distributions. It has been shown—at least for tenants—that the high-income groups are likely to be underrepresented and the very poor groups, overrepresented. The net effect of the two opposite biases is uncertain on *a priori* grounds. Applying the same indirect test as for 1933, we compare the degrees of inequality in the Census and Financial Survey distributions of rent for 1929. In 11 cities, inequality shown by the Census data exceeds that found for the Financial Survey distributions, while the opposite tendency prevails in the other 21 cases.¹⁹ In comparison with the total, the 1929 rent samples thus show too high rather than too low a degree of inequality. It may be expected, therefore, that in most cities the sample *income* distributions for 1929 tend to overrate the degree of income inequality for that year.

Inequality in the rent distribution increases from 1929 and 1933 in most cases, whether we consider the Financial Survey sample or the total population figures. Among the Survey samples are 6 exceptions; among the 100 per cent samples of the Census and Real Property Inventory, 2. Therefore the increase in *income* inequality during this period observed in the samples cannot well be ascribed to deficiencies in the coverage of the usable samples. On the contrary, the observed deficiencies tend to produce an *apparent decrease* in inequality from 1929 to 1933—except for Providence, where they might account for an increase.

To test whether these deficiencies might explain the observed decrease in income inequality *among the upper income recipients*, the change in the degree of *rent* inequality was studied for the higher rent groups, in one case with the help of the Financial Survey data, in the other case with the Census and Real Property Inventory material. In 20 cities, for which the proper data are available, the rent distributions were subdivided at the same cumulative percentage of families as the corresponding income distributions. In contrast to the result obtained

¹⁸ Similar tests, carried out for owners using the 1933 Financial Survey and Real Property Inventory value distributions, give the same result: in all cities, the Real Property Inventory distributions show higher dispersion of values than the Financial Survey samples.

¹⁹ No adequate Census data available for Boise.

TABLE A 2
Coefficients of Concentration, Distribution of Tenants by Rent
Usable Samples

	1929		1933	
	FINANCIAL	CENSUS	FINANCIAL	REAL
	SURVEY	(2)	SURVEY	PROPERTY
	(1)		(3)	INVENTORY
	(1)	(2)	(3)	(4)
Atlanta	.384	.380	.370	.383
Birmingham	.379	.372	.345	.321
Boise	.245	*	.263	.294
Butte	.263	.262	.281	.298
Cleveland	.247	.223	.263	.273
Dallas	.246	.280	.253	.302
Des Moines	.267	.281	.278	.301
Erie	.219	.214	.235	.242
Indianapolis	.278	.275	.318	.337
Lansing	.227	.184	.280	.289
Lincoln	.267	.248	.271	.277
Little Rock	.340	.323	.325	.332
Minneapolis	.244	.243	.251	.258
Oklahoma City	.254	.250	.274	.315
Peoria	.262	.256	.269	.285
Portland, Me.	.230	.240	.229	.246
Portland, Ore.	.268	.252	.276	.294
Providence	.252	.273	.256	.261
Racine	.216	.196	.262	.283
Richmond	.361	.344	.348	.355
Sacramento	.215	.222	.232	.245
St. Joseph	.274	.289	.294	.308
St. Paul	.250	.254	.246	.263
Salt Lake City	.266	.272	.285	.295
San Diego	.221	.243	.228	.268
Seattle	.259	.248	.269	.292
Springfield, Mo.	.300	.312	.308	.328
Syracuse	.222	.203	.236	.245
Topeka	.307	.303	.312	.320
Trenton	.235	.238	.244	.272
Wheeling	.284	.280	.287	.319
Wichita	.274	.251	.305	.328
Worcester	.220	.218	.235	.243

* Adequate data not available.

In computing these coefficients the mean rent in each of the 5 central rent classes was fixed at the class midpoint. In the case of the 2 extreme classes the mean rents for tenants, Minneapolis, 1933 (Financial Survey) were used. There are no data on mean rents in the sources for columns 1, 2, and 4.

for the income distributions, inequality within the upper rent groups increases in a majority of cases. In the Financial Survey material, it increases in 14 and decreases in 6 cities; in the Census-Real Property Inventory material, it increases in 19 cities and decreases in 1. The fewer increases in the Financial Survey material as well as the tendency

for them to be numerically smaller than those in the Census-Real Property Inventory data may suggest the presence of the suspected bias; but the divergent trends of change in the upper branches of the income and rent distributions of the Financial Survey cast doubt on the practical significance of this test. The upper branches of the income and rent distributions seem to follow different patterns of change. There is no conclusive evidence that the observed decline in income inequality within the upper group can be ascribed to the bias in the Financial Survey material.

The importance of the biases established in the preceding analysis should not be overrated. Even when the bias is definite it does not entirely distort the picture of the distributions. The distributions of owners in San Diego, by value of property (one of the cases in which the

TABLE A 3
 Percentage Distributions of Owner-occupants in San Diego by
 Value of Property, January 1, 1934

VALUE CLASSES (dollars)	REAL PROPERTY INVENTORY	FINANCIAL SURVEY (usable sample)
Under 1,000	7.6	2.9
1,000- 1,499	7.5	5.6
1,500- 1,999	10.4	9.0
2,000- 2,999	25.0	25.4
3,000- 4,999	29.4	37.4
5,000- 7,499	12.6	12.6
7,500- 9,999	3.3	3.4
10,000-14,999	2.1	2.3
15,000 & over	2.0	1.4
Total	100.0	100.0

Financial Survey distributions deviate most from the Real Property Inventory) may serve as an example. The discrepancies are very noticeable, for the lower value groups but tolerably small for the upper groups. (see Table A 3).

The Financial Survey income distributions show a plausible relation to the distributions established by the Study of Consumer Purchases for 1935-36. Both surveys covered Atlanta, Butte, Portland, Oregon, Providence, and Springfield.²⁰ Since 1935-36 is in a period of recovery we should expect the income distribution to occupy some intermediate

²⁰ In Atlanta and Providence the two surveys did not cover the same area. The Study of Consumer Purchases sampled the city proper, while the Financial Survey sampled the entire metropolitan district. The difference in coverage is more serious for Providence. According to the Real Property Inventory of 1934, 64,109 families lived in the city proper of Providence and 164,977 in the metropolitan district; 74,621 in the city proper of Atlanta and 98,593 in the metropolitan district.

position between the extremes of 1929 and 1933. In a general way, this is confirmed by the comparison of the data in Table A 4. In the lower income groups, the frequencies of the Financial Survey for 1929 tend to fall below, those for 1933, above, the frequencies of the Consumer Purchases Study. In the higher income groups, the relationship is reversed, 1929 showing higher, 1933 lower, frequencies (except in Providence). However, in four cities the frequencies in the \$1,000-1,499 bracket are higher for 1935-36 than for the two extreme years, and in two cities, the same is true for the \$500-999 bracket. The median family incomes for 1935-36 are between those for 1929 and 1933.

TABLE A 4
 Percentage Distributions of Families and Median Incomes
 Financial Survey and Consumer Purchases Study

FAMILY INCOME (dollars)	ATLANTA, GA.		BUTTE, MONT.		PORTLAND, ORE.		PROVIDENCE, R. I.		SPRINGFIELD, MO.				
	Financial Survey identical sample 1929 1933	Con- sumer Pur- chases Study ¹ 1935-36	Financial Survey identical sample 1929 1933	Con- sumer Pur- chases Study ² 1935-36	Financial Survey identical sample 1929 1933	Con- sumer Pur- chases Study ³ 1935-36	Financial Survey identical sample 1929 1933	Con- sumer Pur- chases Study ⁴ 1935-36	Financial Survey identical sample 1929 1933	Con- sumer Pur- chases Study ⁵ 1935-36			
Under 500	19.4	34.5	9.1	34.5	13.8	32.4	15.4	10.0	22.0	14.3	16.3	31.1	24.2
500- 999	22.9	21.9	11.3	26.9	16.8	24.8	21.7	16.2	26.3	23.9	19.5	25.7	25.6
1,000-1,499	14.5	13.8	17.8	14.0	21.2	17.4	21.3	23.4	21.6	24.0	22.9	18.5	19.7
1,500-1,999	12.0	11.0	27.6	10.5	18.6	13.2	17.6	19.1	13.3	14.8	17.1	12.6	14.3
2,000-2,999	16.5	10.9	20.9	8.6	19.1	7.8	16.6	18.8	9.9	11.5	15.8	7.6	11.4
3,000 & over	14.7	7.9	13.4	5.5	10.6	4.3	7.4	12.5	6.9	6.5	8.4	4.4	4.8
Median income	1,264	790	1,715	725	1,458	833	1,273	1,510	1,038	1,133	1,310	847	1,004

¹ Department of Labor, Bureau of Labor Statistics, *Study of Consumer Purchases, Urban Series*, Bul. 647, I, 8.

² *Ibid.*, Bul. 649, I, 8, 9.

³ *Ibid.*, Bul. 645, I, 8.

