Volume Title: Studies in Income and Wealth, Volume 3

Volume Author/Editor: Conference on Research in National Income and Wealth

Volume Publisher: NBER

Volume ISBN: 0-870-14158-9

Volume URL: http://www.nber.org/books/unkn39-1

Publication Date: 1939

Chapter Title: American Studies of the Distribution of Wealth and Income by Size

Chapter Author: C. L. Merwin, Jr.

Chapter URL: http://www.nber.org/chapters/c9521

Chapter pages in book: (p. 2 - 94)

Part One

AMERICAN STUDIES OF THE DISTRIBUTION OF WEALTH AND INCOME BY SIZE

C. L. MERWIN, JR.
BUREAU OF FOREIGN AND DOMESTIC COMMERCE
UNITED STATES DEPARTMENT OF COMMERCE

Discussion

SIMON KUZNETS
NATIONAL BUREAU OF ECONOMIC RESEARCH

AMERICAN STUDIES OF THE DISTRIBUTION OF WEALTH AND INCOME BY SIZE

C. L. MERWIN, JR.

THE genetic development of the analysis of wealth and income distribution by size in the United States is not without a cause. This one is tempted to seek in the strands of economic history. The immediate impulse was a Census study by G. K. Holmes and J. S. Lord, entitled *Farms and Homes: Proprietorship and Indebtedness in the United States at the Eleventh Census.* This special study, provided for by an Act of Congress dated February 22, 1892, was the culmination of discussions then raging in legislative halls concerning the concentration of wealth.

The ultimate causes are farther to seek. The rise of industrial trusts provides one clue. Although evidences of industrial integration in the United States appeared as early as 1861 with the cordage industry agreements, the movement did not gain momentum until the last quarter of the century when the Standard Oil trust was formed. By the conclusion of the initiating trust-proper phase of the movement in the 1890's, statisticians had already inaugurated analysis of the distribution of wealth, by size of wealth holding.

Another clue is provided by the trend of wholesale prices. Over the nineteenth century there was a secular decline in wholesale prices which the Civil War inflation merely interrupted. From the currency restabilization in 1871 to the close of the century, prices fell more than a third. The year 1896 represented the all-time low point. Persons enjoying fixed incomes (e.g., recipi-

cuts of property income) stood to profit from falling prices; while persons burdened with fixed charges (e.g., farmers with mortgages) felt the pinch of the price decline.

The mere existence of these positive correlations does not imply a cause-and-effect relationship between the rise of trusts and the decline in prices on the one hand, and analyses of distributions of wealth and income on the other. Yet such concomitance does suggest a relationship between economic history and academic interests, and warrants the presumption that wealth and income distribution analysis was launched to fill a pressing social need, not merely to provide academic jousts for statisticians.

Although the trust movement and price trends have been carefully described and analyzed by scores of investigators, little has been written on the statistical attempt to analyze the problems raised by this economic and social transition. The object of this paper is to delineate, in Sections I and II, the historical strands of wealth and income distribution analysis. These sections are concluded by recapitulations in outline form, which serve to emphasize the salient characteristics of these earlier studies. In a concluding section speculation is ventured concerning possible reasons why distributions of wealth and of income thus far constructed have been relatively inadequate.

I American Studies of the Distribution of Wealth

1 HISTORICAL AND METHODOLOGICAL REVIEW

The stub of a tabular distribution of wealth, by size, would show a series of wealth classes ranging from, say, '0–$500' to '$1,000,000 and over'. The frequencies would give the number of individuals, families, or some other wealth-holding unit in each class; for example, the number of persons possessing wealth valued at '0–$500' and at '$1,000,000 and over'.

The two substantive elements in the distribution of wealth by size are the nature and dollar amount of the wealth that is distributed, and the nature and number of the wealth-holding units. The first is commonly referred to as national wealth, the estimation of which is a problem all its own.[1] The second hinges

[1] See Simon Kuznets, *Studies, Volume Two* (1938), Part One.

on a decision as to among whom (or what) the wealth is distributed. We could, for example, tabulate the distribution of wealth, by size, among individuals, families, estates, corporations, and other more or less homogeneous entities. In addition, the distribution could be by subdivisions of each of these units. In this paper the distribution of wealth is considered with respect to the individual, family, and estate units.

No complete census of wealth holdings by any of these units has ever been taken in the United States. Therefore, attempts to construct a distribution of wealth must rely on samples of the universe, or on wealth's possible functional relationship with some other variable such as income, tax payments, house ownership. When samples are used, there is the problem of extending the partial picture to give a complete description. Frequently the other aids mentioned above are employed in this task, but sometimes the extension of the sample is a matter of sheer guesswork. To enhance its applicability and augment its coverage, the sample may be treated beforehand, by means of supplementary data and arbitrary assumptions. In any case the problems confronting the investigator are numerous and difficult, as the descriptions of these studies on the following pages illustrate.

a) *Holmes' attempt*

At least two publicized attempts were made in the last decade of the nineteenth century to estimate the distribution of wealth in the United States.

The first, by G. K. Holmes in 1893,[2] was a modest statistical inquiry, based on census data, into the number of families of different economic characteristics in the United States and the wealth possessed by each class of family. Of the 12,690,152 families enumerated in the 1890 Census, 11,593,887 were classified into six categories which included farm-hiring families, families owning encumbered farms, families owning free farms, home-hiring families, families owning encumbered homes, and families owning free homes. The allocation of families to these categories was accomplished by a complicated procedure involving farm and home proprietorship data, averages of the farm and home possessions and indebtedness of the various types of fami-

2 'The Concentration of Wealth', *Political Science Quarterly*, VIII (1893), 589–600.

lies, assumptions as to the number of farms and of families occupying non-farm houses, and arbitrary allowances for 'other' possessions and debts of each class of family. Together these families were estimated to possess $17,356,837,343. Since the national wealth was set by Holmes at "about sixty billions of dollars",[3] 91 per cent of the families, therefore, owned 29 per cent of the wealth, and, by subtraction, 9 per cent of the families owned 71 per cent of the wealth. Having estimated the wealth of the poorer class, Holmes directed attention to that of the very rich. According to a *New York Tribune* estimate of 1892,[4] there were 4,047 millionaires in the United States. Holmes assumed that their average wealth was $3,000,000; which meant that they held 20 per cent of the total wealth. His final distribution of wealth, in Lorenz curve form, was:

.03 per cent of families (i.e., the millionaires) own 20 per cent
9 per cent of families (excluding millionaires) own 51 per cent
91 per cent of families own 29 per cent.[5]

From added comments of Holmes concerning the wealth distribution among the poorer classes, it is possible to split up this distribution of wealth in 1890 into five classes, from rich to poor:

PERCENTAGE OF FAMILIES		PERCENTAGE OF WEALTH	
SIMPLE	CUMULATED	SIMPLE	CUMULATED
.03	.03	20	20
8.97	9.00	51	71
27.00	36.00	20	91
12.00	48.00	4	95
52.00	100.00	5	100

b) *Spahr's distribution*

The second pre-twentieth century estimate of the distribution of wealth, statistically more pretentious than the first, was published

[3] *Ibid.*, p. 590. The Census estimate of the total value of tangible property in the United States was $65,000,000,000; see *Compendium of the Eleventh Census: 1890* (Washington, 1898), Part III, p. 94. Holmes' figure seems designed to approximate this estimate, and perhaps the fact that it was made five years earlier explains why it fell five billion dollars short.

[4] This, and a similar *New York World* list of millionaires, are described in G. P. Watkins' 'The Growth of Large Fortunes', *Publications of the American Economic Association*, 3d ser., VIII (1907), 141–7.

[5] Holmes, *op. cit.*, p. 593.

by C. B. Spahr in 1896.[6] He based his analysis on figures for probated estates obtained from the Surrogate records of New York State. Data were collected for 36 counties, including the area comprised by New York City and Brooklyn, and having a population of 4,625,000 persons, for October, November, and December of 1892. Because they were not deemed representative, the figures for New York City and Brooklyn were excluded, leaving the accompanying distribution of probated estates, which was used as the basis for the subsequent distribution of wealth in the United States.[7] Once these basic data were acquired, generalized

WEALTH CLASS	ESTATES	PERCENTAGE OF ESTATES	REALTY	PERSONALTY	TOTAL WEALTH	PERCENTAGE OF WEALTH
$50,000 and over	36	2	$2,188,540	$6,606,123	$8,794,663	55
50,000–5,000	409	22	2,950,325	2,233,871	5,184,196	32
Under $5,000	1,427	76	989,668	1,095,130	2,085,098	13

assumptions and personal observation ("common observation shows") were relied upon to effect the transmutation of this distribution for certain New York counties into one for the entire country. Spahr reasoned that the figure in this category should be increased about one-half to allow for the many small real estate holdings not recorded in rural counties. Similarly, large personalties were underestimated to avoid the tax, and small ones were eaten up to pay debts; so the latter should be cut one-half. Effecting these transformations,[8] he arrived at a 'corrected' distribution of these New York estates.[9] Thus far it has been pos-

WEALTH CLASS	ESTATES	PERCENTAGE OF ESTATES	REALTY	PERSONALTY	TOTAL WEALTH	PERCENTAGE OF WEALTH
			(millions of dollars)			
$50,000 and over	36	2	2.25	6.75	9	56
50,000–5,000	409	22	3.00	2.00	5	31
Under $5,000	1,427	76	1.50	.50	2	13

sible to follow Spahr's statistical juggling even though one may disagree with certain of his assumptions; but in the transforma-

[6] *The Present Distribution of Wealth in the United States* (New York, 1896).
[7] *Ibid.*, p. 64. Spahr did not compute the percentages for his distributions, but since they are utilized in the argument, they are inserted in the tables.
[8] And rounding off the resulting figures to the nearest quarter million. Spahr could well have added.
[9] *Ibid.*, p. 65.

tion of this last distribution into one for the entire country, even the statistical manipulations are hard to perceive. He had announced his intention of "applying these proportions [of the above table] to the nation at large",[10] but he modified this resolve by saying "with much precision" that one-eighth of the "families"[11] of the country hold property worth more than $5,000. This decision was based on the distribution of estates in New York City, the Census investigation of farm mortgages, the distribution for New York State outside the two large cities and the assumption of "a normal death-rate".[12] The further division of this one-eighth between '$50,000 and over' and '$50,000–5,000' was apparently harmonized with the proportion (2 to 22) exhibited in the distribution above, so that the table "for the nation at large" becomes: [13]

WEALTH CLASS	FAMILIES (*thousands*)	PERCENTAGE OF FAMILIES	AGGREGATE WEALTH (*billions of dollars*)	PERCENTAGE OF WEALTH
$50,000 and over	125	1	33	51
50,000–5,000	1,375	11	23	35
Under $5,000	11,000	88	9	14

Even if we accept as sufficiently justified Spahr's division of family holdings into 12 per cent over and 88 per cent under $5,000, there is still the question how he distributed aggregate wealth. If these proportions were meant to follow those in either of the preceding tables, then his arithmetic was 'rough' in the direction of decreasing the inequality of wealth distribution.[14]

[10] *Ibid.*, p. 64.

[11] Spahr changed his terminology from 'estates' to 'families' without warning or explanation. In the rest of his analysis he seems to use 'families' and 'estates' almost indiscriminately. Yet by a family he tells us (p. 66n) that he means "a family of five".

[12] *Ibid.*, p. 66.

[13] *Ibid.* Spahr states (p. 66n) that "nearly one billion dollars [has been] added [to the aggregate wealth of the 'under $5,000' category] for small estates containing only household goods and the like". It will be observed that the total aggregate wealth, $65,000,000,000, is that given by the Eleventh Census for the true valuation of the tangible property in the United States (see footnote 3 above) and the total number of families is approximately that given by the same Census (*Compendium*, Part I, p. 856).

[14] On the other hand, the changes he made in the percentage of families in each class served to increase the inequality of wealth distribution relative to that in the preceding tables.

Spahr did not elucidate this transition, but went on to subdivide the 'under $5,000' class into '$5,000–500' and 'under $500' categories. The Census returns indicated that in the cities the number of families owning over $500 worth of property was 'perhaps' one-third greater than the number owning their homes, while in the small towns and rural districts it was 'perhaps' one-sixth greater. As few holdings of real estate were valued at less than $500, "in the nation at large" the families worth more than $500 numbered 'perhaps' 1,000,000 more than those that owned their homes or farms. That is, about 7,000,000 were property-owning and about 5,500,000 could "justly be spoken of as propertyless". Under an assumption that the latter, as a rule, had household property worth $150, Spahr's final distribution of wealth for 1890 stood as follows: [15]

		PERCENTAGE OF FAMILIES		AGGREGATE	PERCENTAGE OF WEALTH		
WEALTH CLASS	FAMILIES *(thousands)*	SIMPLE	CUMULATIVE	WEALTH *(billions)*	SIMPLE	CUMULATIVE	FAMILY AVERAGE
$50,000 and over	125	1	1	$33.0	51	51	$264,000
50,000–5,000	1,375	11	12	23.0	35	86	16,000
5,000– 500	5,500	44	56	8.2	13	99	1,500
Under $500	5,500	44	100	.8	1	100	150

Before continuing the historical summary, it is interesting to compare the wealth distributions for 1890 constructed by Holmes and Spahr. Chart 1 shows these two independent estimates in the form of Lorenz curves. The percentages of wealth are plotted along the X-axis and the percentages of families along the Y-axis. Both sets of percentages are cumulated, from rich to poor. The reference points are meagre, and the straight lines connecting them are merely aids to the eye, not indicators of where the intermediate points would fall. The difference in the inequality indicated by the two curves is significant, but not so striking as one might have expected considering the dissimilar methods and the many arbitrary assumptions of the two investigators. The greater inequality shown by Spahr's curve is probably largely attributable to the nominal value he placed upon unreported estates, and to certain other statistical juggling in which he indulged. It is hard to say which distribution is closer to the actual distribution of wealth.

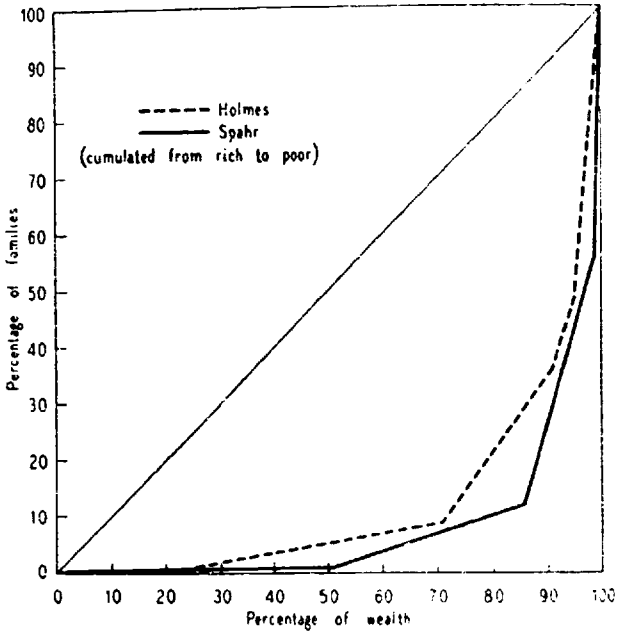
[15] *Ibid.*, pp. 68, 69.

c) *Contributions of W. I. King*

A decade elapsed before another inquirer seriously attempted a distribution of wealth even for selected sections of the country, and another twenty years before a third attempt was made to distribute, by size of holding, the nation's material wealth.

Chart 1

LORENZ CURVES OF HOLMES' AND SPAHR'S DISTRIBUTIONS OF WEALTH, UNITED STATES, 1890



The pioneer work in the field of the distribution of wealth and income by size in the United States was done by W. I. King in 1915.[16] Although he did not venture to derive a complete distribution of wealth, his familiarity with statistical tools makes his analysis of the Massachusetts probated estates data sufficiently important to warrant mention in this survey. The original data, themselves a landmark, are contained in the *Twenty-Fifth Annual Report* (1894) of the Massachusetts Bureau of Statistics of

[16] *Wealth and Income of the People of the United States* (New York, first published 1915, printing cited here is that of 1923), especially pp. 66–76.

Labor, and comprised the values of estates probated in Massachusetts during the four triennial periods 1829–31, 1859–61, 1879–81, and 1889–91.[17] The estates were classified as to ownership by males or females. For 40 per cent of the estates no inventory was filed. King excluded the estates of females and assumed that the non-inventoried estates were of the same size and distribution as those filed with inventories. He found from Census reports that the number of deaths of males 25 years or over in Massachusetts for the three periods considered (1859–61, 1879–81 and 1889–91) exceeded the number of estates filed. He assumed that these non-probated estates were insignificant in value, with an upper limit of $500 and an average value in the first period of $375 and in the other two periods of $400. The resulting distribution contains twelve categories ranging from $0 to $500,000. A similar analysis was made of estates probated during 1900 in six Wisconsin counties, the original data for which appeared in an unpublished manuscript by M. O. Lorenz. No attempt was made to derive from these Massachusetts and Wisconsin data a distribution of wealth for the entire country.

When King returned, some twenty years later, to the task of constructing a distribution of wealth,[18] his insight into the prob-

[17] C. D. Wright left the Massachusetts Bureau to head the new National Labor Commission in 1888, but he was nonetheless instrumental in launching this survey begun "some years" before publication of the preliminary results in 1894 (Massachusetts Bureau of Statistics of Labor, *Twenty-fifth Annual Report*, 1894, p. 55). G. K. Holmes also assisted the Bureau in this work.

[18] W. I. King, 'Wealth Distribution in the Continental United States at the Close of 1921', *Journal of the American Statistical Association*, XXII (1927), 135–53. This article represents the product of a much more extensive investigation than its length would indicate. King became associated with the National Bureau of Economic Research soon after its establishment in 1920, and continued the study of wealth and income distributions initiated in his first book, *Wealth and Income*. With a corps of assistants he constructed distributions of both wealth and income for the United States in 1921. As a result of his labors, two book-length manuscripts now on file at the National Bureau of Economic Research were prepared. One, entitled 'The Distribution of Earnings, Income and Wealth in 1921', never progressed beyond the typewritten stage, although it was completed and signed by King on October 1, 1925. The other, entitled 'Gradations of Earnings and Income in 1921', apparently came nearer publication, for it was mimeographed and given a table of contents and a title page with a 1926 dateline. The latter manuscript was a recasting of the first half of the former manuscript.

Curiously enough, the article here cited was essentially an abstract of the second

lem had broadened considerably. He not only realized that the distribution of wealth among decedents was far from being the distribution of wealth among the living, but he even conceded the criticism of Judge R. S. Galer and W. R. Ingalls that the distribution of wealth among decedents also did not measure the distribution of wealth among persons near the end of their careers. The latter relation, previously claimed by King, was challenged on the grounds that (a) many estates are not probated at all, (b) some property is held by joint title so that no court record is necessary on the death of one of the title-holders, (c) some property is transferred at death without any record of its value (i.e., it is not inventoried), (d) gifts often anticipate death. As a consequence, King concluded that the distribution of wealth may be approached through three channels: distribution of (a) estates, (b) wealth among persons shortly before death, (c) wealth among all the inhabitants of an area.[19] Since no data were available on the second type of distribution, King was limited to estimating the distribution of wealth by the first and third approaches.

The Federal Trade Commission in its study of *National Wealth and Income*, published in 1926, presented data on estates probated during 1912–23 in twenty-four counties in twelve widely scattered states and the District of Columbia. By estimating from Census reports the number of wealth-owners (defined to be 'gainfully employed') who died in these counties during these years, and by assigning to unreported estates an arbitrary value of $100, King constructed an estimate of the distribution of wealth among decedents.[20]

portion of the former typewritten manuscript which, as just noted, apparently did not come as near publication as the first portion on income.

The present investigator was given permission by the National Bureau of Economic Research to read both these manuscripts. This makes it possible, in this section, to amplify the description of King's methods; and in the next section, to describe a hitherto unpublicized distribution of income. In the discussion of King's 1921 distribution of wealth, reference will be made to the published article rather than to the unpublished manuscript, wherever possible.

[19] King, 'Wealth Distribution . . .', pp. 141, 143, 144. King refers specifically to Ch. X of W. R. Ingalls, *Current Economic Affairs* (York, Pa., 1924). On p. 144 of Ingalls' opus appears a recantation, by King, of the probated estates method.

[20] King, pp. 141, 144, 145. The data King used are contained in Ch. II (especially Table 10 on p. 58) of the Federal Trade Commission report entitled *National Wealth and Income*, Sen. Doc. 126, 69th Cong., 1st Sess. (Washington, 1926). The

In the manuscript the attainment of this objective was more fully explained. Nine-tenths of the $260 billion of wealth in the United States,[21] i.e., $230 billion, were assigned to adults, both male and female. Of these, 4,580,000 died during 1916–21, and to these decedents were allotted $30 billion of the $230 billion of wealth. By extrapolating relevant Massachusetts data, King estimated that two-thirds, i.e., $20 billion, belonged to the 2,420,000 adult males dying during this period. With this as a background, he proceeded along two routes toward the distribution goal. In the first, he plotted on double logarithmic paper the Massachusetts and the federal estates data. Observing that the two curves were parallel in the upper wealth class brackets, he extrapolated the federal data to the lower wealth classes in the manner indicated by the Massachusetts data. The insufficiency of this method became apparent when the total wealth thus distributed was summated: it turned out to be only a third of the previously ascertained total of $20 billion. So this approach was discarded in favor of another.

The second route to the distribution goal was rather more devious. The federal data were first reduced from a gross to a net estate basis, and the class limits correspondingly scaled. Then King proceeded to distribute, by several estate classes over $50,000 and one class under $50,000, the number of estates of adult male decedents, and their values. The federal data distributed the adult decedents that were in the classes over $50,000. The rest of the 4,580,000 who died were put in the 'under $50,000' category. To this distribution were applied the 1890 Massachusetts figures for the percentage of estates belonging to males, the 'under $50,000' class again being the residual. The resulting distribution was reduced to percentages, cumulated, and converted to logarithms. The values of estates were distributed by a similar procedure: those of males alone were made to total $20 billion, and 1890 Massachusetts percentages were used to derive the value of estates of males from the value of estates of both

Federal Trade Commission in its tabulation allotted $258 (the average value of the poorest class of estates probated) to the non-probated estates, while King, as we have seen, allotted only $100 to such estates.

[21] An estimate offered by King on p. 322 of an article entitled 'The Net Volume of Saving in the United States', *Journal of the American Statistical Association*, XVIII (1922), 305–23. This figure is apparently an average value for 1916–19.

males and females. As before, the distribution above the $50,000 mark followed the federal data, while the 'under $50,000' class absorbed what remained.

With both these distributions reduced to logarithms, the next step was to plot them, after which readings were taken from the curve to show the distribution of estates among the various percentages of the holders. This made possible construction of a Lorenz curve and comparison with the Massachusetts data for 1889–91. On the basis of this comparison wealth seemed to have become distributed much more evenly between 1889–91 and 1916–21; so much more, in fact, that doubt was cast upon the reliability of one or both sets of data. After consideration of possible sources of bias in the two sets, King concluded that the true curve probably lay between the line representing the Massachusetts data and that representing the federal estates data.[22]

Developing W. R. Ingalls' method of analyzing inventories and capitalizing income, King constructed the third type of distribution of wealth: among all the inhabitants of the United States.[23] His method was very complicated and the published explanation is meagre. Examination of the manuscript, however, makes possible the following more detailed description.

King's general approach was (1) to distribute the farm wealth among farm owners and tenants, (2) to distribute the non-farm wealth among non-farmers, (3) to combine these distributions into a distribution of wealth among all property-owners.

Net wealth of farm owners was estimated from census records by a complicated system involving sundry assumptions concerning the proportion of agricultural debt borne by farm owners, the proportion of tenants' equipment they possessed, and the

[22] W. L. Crum has subjected these federal estate tax data to rigorous statistical analysis in 'The Distribution of Wealth', *Harvard Business Report*, No. 13 (October 1935). He does not venture a complete distribution of wealth; instead, he ignores the lower wealth classes and analyzes the tail of the distribution, along the lines laid down by Pareto.

[23] King, 'Wealth Distribution . . .', pp. 146–53. See especially Ingalls, *op. cit.*, Ch. X, cited by King. This chapter is a reprint of an article appearing in *Iron Age*, October 4, 1923, which was written to disprove the popular belief that 2 per cent of the people own 65 per cent of the wealth in the United States. This belief, incidentally, is traceable to King's *Wealth and Income*. Although he made no attempt to construct a distribution of wealth, Ingalls concludes that the richest 2 per cent own about one-third of the wealth.

like. The final figure, for the end of 1921, was set at $46.5 billion, which was then parceled out among the 3,928,000 farm owners. This total wealth of farm owners was first distributed by size of farm, on the basis of Census data on the value of farm property in farms having various acreages. Since the Census gave also the number of farms in each size class, the assumption that the proportion of tenant-owned farms in each size class was the same reduced this raw Census distribution to a farm owner basis. The wealth in each size class was then split between owners and tenants in the same proportion as total acreage in each size class was split between these two groups. The final step was to cumulate the wealth (size) classes and farm owner frequencies, and read off at the desired wealth class intervals the corresponding frequencies. Decumulation gave the distribution of wealth among farm owners. The resulting curve was smoothed, and the total wealth made to equal $46.5 billion. By similar statistical procedures and arbitrary assumptions the net wealth of farm tenants was estimated and distributed.

King next turned the spotlight on the distribution of wealth among non-farmers, including agricultural laborers. The first step was to calculate the distribution of holdings in the stocks of corporations. The next was to estimate the corporate bond holdings of each wealth class. The funded debt held by individuals was distributed among income groups in the same proportions as interest payments. The third step was to distribute the holdings of government bonds among non-farmers, which was also done on the basis of interest payments.

The sum of the wealth thus far accounted for—wealth of farm owners, wealth of farm tenants, and securities held by non-farmers—totaled only one-half of the Census estimate of $298.4 billion of privately owned wealth in the United States at the end of 1922. This Census estimate, when adjusted to December 31, 1921 conditions, became $281.2 billion, which agreed fairly well with an independent National Bureau of Economic Research estimate of $291.1 billion. Diverse methods were employed to distribute the other half of the total wealth. Real estate was distributed along the lines indicated by *Statistics of Income* data on "profits from sales of real estate, stocks, bonds, etc.", "rents and royalties", and "interest and investment income"; urban

owner-occupied houses and other consumption goods were distributed according to the current money income received by the corresponding sections of the population; and the value of residual, miscellaneous wealth items was distributed on the basis of the *Statistics of Income* data on "profits from sales of real estate, stocks, bonds, etc.", from "business" and from "partnerships, fiduciaries, etc."

As a result of this manipulation, King succeeded in distributing wealth among non-farmers by income classes. The next step was to pass to wealth classes. The technique, called Method H, was frequently employed by King and merits quotation:

> *"Method of Constructing a Frequency Table from a Table Giving the Total Wealth and Number of Persons in Each of a Number of Irregular Classes*
>
> 1. Cumulate the number of persons. Cumulate the amounts of wealth. Plot the cumulated quantities against each other. Run a smooth curve through the points.
>
> 2. Take *frequent* readings from the curve showing the cumulative numbers of persons and their cumulative wealth at each point.
>
> 3. Decumulate the record showing the numbers of persons to find the numbers of persons in the new classes. Decumulate the wealth readings to find the total wealth in each of the new classes. Divide the wealth in each class by the number of persons in the class to find the average wealth of the class.
>
> 4. Take the mid-points between the cumulative frequencies found in (2), and plot against the average wealth in each class.
>
> 5. Take readings on this cumulative curve at the desired class limits for wealth. Decumulate to find the numbers of persons in each class.
>
> 6. Get an approximate verification of the results by multiplying the mid-point of each class [24] by the average wealth of the class and summating the products. The total should correspond with the known aggregate of wealth.
>
> 7. If it does not approximately correspond, the number of classes in (2) is not large enough. By summating the wealth in separate sections of the distribution and comparing with the decumulated

[24] This appears to be a typographical error in the original manuscript. Presumably 'number of persons in each class' should be substituted for 'mid-point of each class'.

figures in the early part of the curve, it may be possible to locate the region in which the major errors occur. In these regions, more readings should be taken in (2) and the later steps should be repeated. This process should be continued until the results are satisfactory."

The final step in King's construction was to combine the three distributions of wealth among farm owners, farm tenants, and non-farmers. The resulting distribution gave the number of wealth owners, i.e., income recipients, in each of 48 wealth classes ranging from "$0 up to $200" to "$40,000,000 and over". In this manner $281 billion in wealth was distributed among 41 million wealth holders (i.e., income recipients).

Salient features of the inequality in the distribution of wealth were pointed out at various places in the manuscript, by means of simple percentages and Lorenz curves. The latter showed, incidentally, the distribution of Massachusetts estates to be the most unequal of the three distributions, while the distribution of the estates reporting under the United States inheritance tax was the least unequal, and the distribution among the living occupied a middle ground. No conclusions respecting the social desirability of the existing distribution were essayed.

King is credited also with a distribution of wealth for 1928, constructed for the Hanover Bank and Trust Company. W. Tresckow, vice president of the bank, published it in 1931 [25] under the title, 'Estimated Cumulative Distribution of Private Property of Individuals among the Entire Population'. The cumulation is from rich to poor. The distribution applies to the continental United States, as of the end of 1928, and contains forty wealth classes. By means of Lorenz curves it is compared with King's distribution for 1921. No comments concerning the methods or data used in constructing the 1928 distribution are offered by Mr. Tresckow. His sole concern is with the significance of these data for trust departments of banks. Moreover, there seems to be no publication by King describing this distribution.[26]

[25] 'Trust Business Possibilities; The Distribution of the Wealth of the United States and Potential Trustors', *Burrough's Clearing House*, September 1931, pp. 13–15, 43, 44.

[26] In a letter to the writer dated April 4, 1938, King stated that the method was fundamentally the same as that used in calculating the 1921 distribution of wealth.

d) *Doane's 'greater diffusion'*

Since King's endeavors, only one attempt to distribute by size the wealth of the people of the United States seems to have been published: that by R. R. Doane in 1935. In a series of articles in the *Annalist*,[27] Mr. Doane patently set out to justify the present distribution of wealth, using as his basic data probated estates figures previously analyzed by other students of the problem and tax payments information appearing in official publications. The latter procedure holds special interest for us, since by means of total tax payments and certain other information in the Treasury Department's *Statistics of Income* and the Census publication, *Financial Statistics of State and Local Governments, 1932*, Doane constructed a distribution of gross private wealth holdings by income classes in the United States for 1932.[28] He launched his

[27] 'Summary of the Evidence on the National Wealth and Its Increasing Diffusion', July 26, 1935, pp. 115–8;
'An Accurate National Wealth Census: Statistical and Other Limitations', Aug. 2, 1935, p. 158;
'Tax Payments as an Aid to More Exact Measurement of Wealth Distributions', Aug. 9, 1935, pp. 189, 214;
'Changes in the Distribution of Wealth Since 1880: Greater Diffusion Shown', Aug. 16, 1935, pp. 222–4;
'The Geographic Distribution of the Physical Wealth in the United States', Nov. 15, 1935, pp. 676–9;
'Property Ownership by States; Security Holdings, Insurance Equities, etc.', Dec. 20, 1935, pp. 844–6;
'The Division of the National Wealth between Farm and Non-Farm Property', Jan. 31, 1936, pp. 196, 197;
'Distribution of Corporate, Individual and Public Debts and Equities, 1932', May 15, 1936, pp. 718, 719, 725.
Several other *Annalist* articles, not originally intended to be a part of this series, nevertheless belong there:
R. H. Jackson, 'Full Text of Memorandum on the National Wealth and Its Distribution', Aug. 30, 1935, p. 292 (a criticism of the methods and figures used by Doane in his third article);
R. R. Doane, 'Rejoinder', Aug. 30, 1935, pp. 292, 293, 312;
S. N. Whitney, 'Weakness of Data Supporting Conclusion of Increase in Diffusion of Wealth', March 6, 1936, pp. 368, 369, 392;
R. R. Doane, 'Statistical Bases for National Wealth Estimates', March 27, 1936, p. 478 (a reply to Whitney's criticism);
S. N. Whitney, 'Statistical Bases for National Wealth Estimates', April 10, 1936, pp. 542, 562 (a further rebuttal to Doane, in letter form).

[28] See p. 189 of his third article listed above. Incidentally, Doane in several instances cites p. 68 of the Census report on *Financial Statistics of State and Local*

construction by distributing total tax payments (other than federal income taxes) by income classes above $5,000, adding corporate taxes to this total and allotting the rest of the tax bill (ascertained in an unexplained manner) to all income classes under $5,000. Although some of the arithmetic is not clear, Doane seems to have distributed the Census total for annual valuation of property among (1) income classes over $5,000, (2) corporations, (3) a 'non-reporting' group later assumed equal to income classes under $5,000. This allocation was carried out roughly according to a Census estimate that the average tax rate per $100 of assessed valuation was $3.08 in 1932. The tax payments were distributed among the income classes over $5,000 apparently in proportions derivable from *Statistics of Income* data for 1932. His references to this source are too general to allow checking these percentages. Once the general property was distributed, the corporate holdings were dropped out, the group 'non-reporting' was labeled 'under $5,000', and the addition to this distribution of intangible property was undertaken. Relying primarily on *Statistics of Income* data he allocated to the various income classes tax exempt securities, other bonds, notes and mortgages, capital stock, savings and other deposits, and life insurance equities.[29] His resulting distribution presented total gross holdings by income classes, with incomes above $5,000 divided into nine categories, and those under $5,000 included in one category. No figures were given for the number of wealth holders (or of income recipients). During the week a significant transformation of this distribution took place, for in the next (the fourth) article it was summarized in such a fashion that the incomes under $5,000 fell into four classes, and the percentages of total number and value were given not only for each of these four classes but also for each of eight classes over $5,000. Neither the method of ascertaining and distributing the number of wealth holders nor the manner in which the wealth holdings of the 'under $5,000' class were divided into four sub-categories is indicated. No absolute figures are given in the final distribution, only the percent-

Governments, 1932, when he must mean p. 66; a table of contents appears on p. 68.

[29] It was with this phase of his analysis that R. H. Jackson, then Counsel for the Bureau of Internal Revenue, raised his most serious objections. Jackson characterized Doane's figures as "very misleading"; see *Annalist*, Aug. 30, 1935, p. 292.

ages of an unknown total. Apparently it is to be taken on faith, and in any case it is in terms of wealth per income class, not per class of wealth holders.

The rest of Doane's analysis, in which he tries to demonstrate an increasing diffusion or lessening inequality in wealth distribution since 1880, is not of particular interest to us because he uses (sometimes in misleading form) [30] data, prepared by other investigators, with which we are already familiar: Massachusetts estates data for 1879–81, Lorenz's data for six Wisconsin counties in 1900, King's computation of a complete distribution of wealth for the continental United States in 1921, and, finally, his own figures for the distribution of wealth in 1932.

e) *Lehmann's novel method*

In recent years an ingenious method for estimating the amount of wealth held by the richer classes has been employed by Fritz Lehmann. In his contribution to *Political and Economic Democracy* a general outline of the method is presented. In a later publication, it is explained further.[31]

Briefly, the method runs as follows: From the estate tax tabulation in *Statistics of Income* ascertain the average value of estates in each estate class by subtracting 90 per cent of the 'Debts, unpaid mortgages, etc.' from the 'Total gross estate' and dividing the remainder by the number of returns in the given estate class. Determine by correlation the function relating this average value of estate to the item 'Capital stock in corporations',

[30] For example, estates of females were not excluded from the Massachusetts data, although King was careful to subtract them because it could not be expected that their estates would be comparable to those filed by males. By including the estates of females, Doane increased the inequality of his earliest distribution, which had the effect of indicating an increasing diffusion of wealth through time when it was compared with the later figures.

[31] Fritz Lehmann, 'The Distribution of Wealth', *Political and Economic Democracy*, ed. by Max Ascoli and Fritz Lehmann (New York, 1937), pp. 159–75; Gerhard Colm and Fritz Lehmann, *Economic Consequences of Recent American Tax Policy*, supplement 1, (1938) to *Social Research*. See especially pp. 43–53, and Appendix A, prepared by Charles Stewart, entitled 'Method of Estimating the Influence of the Personal Income, Gift and Estate Taxes upon Savings and the Distribution of Wealth', pp. 91–8. For a detailed statement of the technique and evaluation of its advantages and limitations see Charles Stewart, Part Two, discussion by W. L. Crum, Milton Friedman, and Fritz Lehmann, and Mr. Stewart's reply.

found in the same estate tax table. By means of (a) this function, (b) the personal income tax tabulation of 'Dividends on stock of domestic corporation', and (c) an assumed average dividend rate for common stock, compute the average size of estate corresponding to the various income classes. That is to say, from the regression line associating stock holdings with average size of estate 'read' the average size of estate corresponding to the capitalized value of common stock dividends.

There are several statistical defects and arbitrary assumptions implicit in this method, as Lehmann is careful to emphasize. Moreover, it gives only the tail of the wealth distribution, and fails to tell anything about the bulk of the wealth holdings. Finally, it shows the wealth holdings by the constituents of income classes, not of wealth classes; so there remains the problem of passing from income to wealth classes. Nevertheless, the results have a fair share of utility, and Lehmann's analysis is an excellent example of those problems the study of which is facilitated by a knowledge of the size distribution of wealth and income.

2 PURPOSE OF THE STUDIES

It was no mere coincidence that the Sherman Anti-Trust Act and the genesis of a more or less intensive study of the distribution of American wealth both occurred in the last decade of the nineteenth century. The first phase of the so-called trust movement—characterized by trusts-proper, such as the original Standard Oil combine of 1879—was drawing to a close, and the growth of monopolies was about to enter upon its second phase, that of holding companies and giant consolidations. Moreover, the secular fall in prices had reached its trough.

a) *Early students ethically motivated*

Both Holmes and Spahr seemed to be gravely concerned about the inequality in the distribution of wealth indicated by their estimates. They were apparently more interested in the social implications of the figures they compiled than in the accuracy and representativeness, from a statistical standpoint, of their resulting distributions. Holmes did not make a specific study of the tax problem in his article, yet he did suggest "progressive taxes on income, gifts and inheritances" to keep the concentra-

tion of wealth from going too far.[32] Spahr, although his book was labeled *The Present Distribution of Wealth in the United States,* nevertheless felt that the inequality in this distribution warranted devoting the concluding portion of his text to the problem of taxation, especially the inequity of the tax burden in relation to the distribution of wealth and income. Singularly enough, although he was writing before the days of our income tax, he concluded that the tax burden with respect to income was relatively just, but with respect to wealth, relatively unjust. He even forecast a progressive property tax, so alarmed was he by the widening gulf between classes. Finally, Spahr pondered taxation as a solution to the wealth distribution problem long enough to perceive that "the future laws which shall make better or worse the distribution of property are likely to accomplish their end, not by the bodily transfer of property from one class to another, but by making more equal or more unequal the distribution of the future incomes of the people".[33] This quotation confirms a suspicion held as early as Spahr's day that the real key to the problem of the concentration of wealth resided ultimately in the distribution of income.

b) *King's purpose statistical*

By the time King made his analysis of the *Wealth and Income of the People of the United States* in 1915, he was able to say without serious danger of being controverted that the distribution of income was more important than the distribution of wealth, and that the latter would not need to be analyzed, were it not that the possession of wealth gives power. Before launching his statistical inquiry, King discussed in general terms the problem of wealth concentration, and concluded that only a moderate (not the existing) inequality in distribution of wealth was justified by social and economic considerations. Not until the end of the book did he revert to the ethical problem involved in wealth and income distribution, when he cited population as a controlling factor, and emphasized the slogan "Poverty must go".[34] No program of taxation was proposed, and a transfer of

[32] Holmes, *op. cit.*, p. 600.
[33] Spahr, *op. cit.*, p. 73.
[34] King, *op. cit.*, pp. 238–55.

wealth was frowned upon. The problem foremost in his mind seemed to be statistical. He was concerned with constructing an accurate and representative distribution of wealth for 1910. It has already been pointed out that this analysis was only for selected sections, not for the country as a whole. Therefore, our chief interest centers on a later work by him, in which a complete distribution of wealth for the continental United States was essayed.

In this second study, King seems to have changed his mind somewhat as to the usefulness of wealth distribution analysis, for he states that "from the social standpoint, nothing can be of greater significance" than the distribution of wealth per person or per family.[35] As before, he asserted that "the outstanding characteristic of wealth is that to its owner it gives power", and that "the possession of wealth is a great convenience".[36] He now emphasized, perhaps more than before, the political significance of wealth, a wide diffusion of wealth being taken to imply political stability. Aside from these brief comments, King in his second study was concerned solely with the statistical problem of constructing a distribution of wealth among the inhabitants of the United States. Even the slight ethical tinge of his preceding study is absent, by design.[37]

c) *Doane an apologist*

The purpose of the most recent complete distribution of wealth is not far to seek. Doane is an apologist for the present concentration of wealth in the United States, and his purpose was not only to show that wealth concentration is decreasing but that the current inequality in the distribution of wealth is justified on the basis of age differences in the population. The problem of statistical analysis seemed to be secondary, though the study

[35] 'Wealth Distribution in the Continental United States at the Close of 1921', p. 139.

[36] *Ibid.*, p. 140.

[37] *Ibid.*, p. 153. No elaboration of this teleological design is offered in the manuscript. King dismissed the question by referring in the Introduction to two groups particularly interested in the distribution of wealth and income: reformers and sales managers. The former need to know the facts about inequality of wealth holdings and income receipts in order the better to carry out their social programs. The latter are anxious to know how wealth is distributed and income divided in order to gauge correctly the demand for their products.

is replete with figures. It is to be expected that Doane would steer clear of such problems as redistribution by taxation, political stability, and social security, which engaged earlier students of the wealth question.

3 STATISTICAL ADEQUACY OF THE STUDIES

So far as statistical adequacy is concerned, all our inquiries have been impeded by a dearth of pertinent data. In addition each study has individual defects.

Holmes, relying on Census data of farm and home proprietorship, did not construct a frequency distribution of wealth holdings; he was content with noting, after the fashion of Lorenz curve analysis, the proportions of wealth held by given proportions of the population, and no rigorous accuracy for these figures was claimed. In general, Holmes' study presents only rough estimates of the general concentration of wealth holdings in the United States, and is not quite in the same class with the later studies.

a) *Spahr's weaknesses*

Spahr, by utilizing a method long popular in Europe, attempted to construct an actual frequency distribution of the wealth holdings for the entire United States. We have seen that he relied on probated estate records for New York State outside New York City and Brooklyn, and on certain Census farm mortgage data respecting the value of farms. In addition to his too free use of 'common observation' when statistics were either few or biased,[38] Spahr's analysis is open to the following objections:

1. It seems improbable that New York State outside of the metropolitan area was representative of the entire country, especially in 1890, with respect to the distribution of wealth. Not

[38] An oft-quoted statement from the Preface (p. v) of Spahr's book, follows: "The conclusions reached respecting the present distribution of property and incomes are in the main those which common observation has forced upon thoughtful men and women in the ordinary walks of life. The writer has learned, and hopes to teach, that, upon matters coming within its field, the common observation of common people is more trustworthy than the statistical investigations of the most unprejudiced experts. Indeed, he has come to believe that social statistics are only trustworthy when they show to the world at large what common observation shows to those personally familiar with the conditions described."

only was New York State one of the first to be colonized and settled, but it was also industrial while the states in the South and growing West were predominantly agricultural. The one link made to farm mortgage data [39] seems insufficient to compensate for this basic dissimilarity.

2. Even if the data for New York State were representative of the entire country, there is still the question of how closely a distribution of wealth among decedents represents the more realistic concept of the distribution of wealth among the 12,500,000 families in the United States. As mentioned above, this defect has long been recognized, for not only are many estates never filed for probate, but some of those which are filed have no inventories attached, considerable property is held jointly (e.g., by husband and wife), and gifts in anticipation of death are common. In addition there is a tendency to underestimate large estates for tax reasons and exaggerate small ones by failing to specify the debts. By adjusting the original data Spahr tried to overcome some of these defects, but not until King's first study was a systematic attempt made to correct for these errors.

3. As already pointed out, 'it is estimated' is the weak point in Spahr's entire analysis, and his resulting distribution of wealth was little more than a guess, bearing only general similarity to the probated estates data originally intended to be basic.

4. An identity was assumed between estates and families that is neither explained nor readily apparent.

5. Finally, no careful definition of wealth was attempted. The concept employed seems to involve both realty and personalty, while the chance that there might be overlapping between the two in his complete distribution was not mentioned. The total aggregate wealth actually allocated to the 12,500,000 families was apparently a Census estimate of the tangible property in the United States.[40]

[39] I.e., his estimate that about one-eighth of the farms seemed to be worth more than $5,000 each.

[40] According to the *Compendium of the Eleventh Census, 1890*, Part III, p. 94. "The true valuation of all tangible property in the United States, exclusive of Alaska, at the close of the Census period, 1890, amounted to $65,037,091,197." No account is taken of "credit money, or of promissory notes, mortgages, or securities, although such items are frequently subject to ad valorem taxation". 'True valuation' is construed to mean 'fair selling price'. Real estate constitutes two-thirds of

b) *King relatively satisfactory*

King, in both his studies, gave evidence of being a relatively thorough and careful statistician. Yet defects are present. In his book, *Wealth and Income,* he attempted partial coverage in his distribution, using probated estates data for Massachusetts and Wisconsin. Such records are open to the objections pointed out in connection with Spahr's study, while King's attempts to overcome some of the more obvious defects in these original data are questionable. He offers no justification for his maximum limit of $500 assigned to non-probated estates of Massachusetts males who died when 25 years or older, and one wonders why he should have assigned an average value of $375 to such estates in the first period studied and $400 in each of the other two. A. A. Young has suggested as a further criticism of King's method that he should have allowed for the much greater inequality of possessions among men at the close of life than among men with a normal age distribution.[41] Finally, even for Massachusetts, the distribution constitutes only a sample, since the 40 per cent of the estates filed without inventories were assumed to be distributed in the same proportions as the other 60 per cent. A similar criticism is applicable to all the distributions based on probated estates records.

King's second, more ambitious attempt to construct a distribution of wealth for the entire country was so inadequately explained in the published article that evaluation of it must have reference to the manuscript description. The use of probated estates data—the method King employed in his first 1921 distribution—has already been criticized. It need only be added that

this $65 billion total, with railroads ranking second. Other items listed are plant machinery and raw materials plus finished goods on hand, farm inventories including livestock, mines and quarries, gold and silver, and communications, shipping and canals. According to Part I of the *Compendium*, p. 856, the number of families is put at 12,690,152.

[41] A. A. Young's review of King's *Wealth and Income*, *Quarterly Journal of Economics*, XXX (1916), 583. This criticism, while not obvious *a priori*, may be borne out by King's 1927 article in which the distribution of wealth among all the people of the United States in 1921 was shown to be less unequal than the distribution of wealth among decedents in twenty-four scattered counties during 1912–23 (p. 151). On the other hand, the method of constructing the 1921 distribution (on the basis of income classes) may have been such as to attenuate the inequality in the distribution.

the divergence among the distributions obtained by the different applications of this method—i.e., Massachusetts data of 1889–91 and King's two wealth distributions based on probated estates and described in the manuscript—does nothing to dispel the doubt cast over the results. The process of capitalizing income—King's second 1921 method—is likewise dangerous because the returns from different but monetarily equal units of capital vary greatly. Since two persons with the same income from capital may have widely different amounts of capital, it would seem that a distribution of wealth constructed by capitalizing income is essentially a distribution of income. Wealth holders are classified by income classes and the aggregate wealth is distributed among these wealth holders roughly in proportion to their incomes. Such results may give a general idea of the distribution of wealth; but as frequency distributions amenable to measurement and interpretation, they are obviously inadequate. In addition, this method required the assumption that the class of wealth holders is identical with the class of income recipients. Unless the wealth tally was sufficiently refined to register relatively minute holdings, it would seem that the latter class was larger than the former. Also, if capital losses were taken into account, it might well be that certain persons with wealth would still have no income. Finally, the assumptions required in utilizing the farm and income tax data were not only numerous but also arbitrary.

c) *Doane confusing*

Compared with Doane's construction of a distribution of wealth for 1932, King's statistical method is a model. Doane's analysis is more heavily documented, but not much more effectively, since certain page references are so general as to be virtually useless. Because Doane follows the principle that the real property of an individual is some multiple of his tax payments, he is open to a criticism similar to that inveighed against King: an average tax bill per unit of assessed valuation is bound to conceal variations that would alter radically the wealth holdings of individuals. Further, his method of passing from assessed valuation to real value is based on another general average derived from National Industrial Conference Board figures, and is open

to the same criticism of concealing significant variations among properties. Since not only assessed valuations but also tax rates may vary markedly from section to section, and among kinds of property within a section, such estimates of the distribution of real property are questionable. Doane's distribution of such personalty as securities, life insurance, and savings deposits has not only been found factually wanting by R. H. Jackson,[42] but also involves the previously criticized principle of estimating wealth by capitalizing income. In general, it seems that Doane greatly exaggerated the number of holders of such personalty by failing to consider duplications arising from the fact that one person may hold stock in several companies, that many life insurance policies are industrial and others weekly (among wage earners, especially), and that a person may have life insurance policies and savings deposits in more than one institution. Moreover, Doane's resulting distribution of percentages explains neither how the number of wealth holders was estimated and distributed nor how the class interval of 'under $5,000' was subdivided into four categories. Finally, Doane does not convert his 'wealth holdings by income classes' into the more consistent 'wealth holdings by wealth classes'. Doane's distribution resolves itself into a distribution of wealth arranged in the proportions in which income is distributed,[43] which is in turn made to follow the distribution of tax payments. Precisely what meaning such a distribution has is hazardous to predict.

In general, the statistical picture presented by these attempts to construct distributions of wealth holdings by size in the United States is as gloomy as the picture of our concentration of wealth itself is to some people. Not only is there a paucity of pertinent data, but (a) no decision has been made as to what constitutes wealth—what, that is, should be distributed among the individuals or families, (b) there is no agreement whether wealth distributions should be on the basis of individuals or families (or

[42] *Annalist*, August 30, 1935, p. 292.

[43] This would deceptively show a greater diffusion through time (when compared with earlier distributions based on estates), for income is distributed more evenly than wealth because human skills and capacities are not included in wealth estimates.

estates),[44] and (c) there is some question as to the intrinsic usefulness of a distribution of wealth, when a distribution of income is contemporaneously available. King suggested as the chief merit of the former that it revealed the distribution of power and of security against emergencies. But it may also be argued that the distribution of income is equally revealing as to the distribution of economic power, and more important in certain tax problems, in analyses of savings and the velocity of money, in the problem of welfare from the subsistence and standard of living viewpoint, and in economic theory. The problem is complex, but it has yet to be proved that a distribution of wealth is of as great intrinsic value in the study of social problems as a distribution of income.

4 RECAPITULATION OF WEALTH DISTRIBUTION STUDIES

The salient characteristics of these earlier studies of the distribution of wealth can perhaps best be contrasted by an outline that emphasizes in summary form major points of similarity and difference.

NAME	OPUS	DATE OF DISTRIBUTION	BASIS OF ESTIMATE	UNIT	PURPOSE, REMARKS
G. K. Holmes	1893	1890	Tangible wealth, Census estimate	Census family	Social—suggested restrictive taxes: only outline of frequency distribution given.
C. B. Spahr	1896	1890	Estates probated in N. Y. State	Estate Family	Social — taxation problem analyzed; perceived income distribution as vital.
W. I. King	1915	1860 1880 1890 1900	Estates probated in Mass., and Lorenz estates data for six Wis. counties	Estate	Statistical. Complete coverage of U. S. not attempted. Income distribution more revealing.

[44] In addition, in those distributions among families, no attempt is made to refine the Census concept of what constitutes a family.

NAME	OPUS	DATE OF DISTRIBUTION	BASIS OF ESTIMATE	UNIT	PURPOSE, REMARKS
W. I. King	1927	1912–1923	Federal Trade Commission (1926 Report) data on estates probated in 24 counties	Estate	Statistical, but links wealth diffusion with political stability. Not complete coverage.
		1921	Capitalize income	Property owners	Converted income classes into wealth classes; complete coverage claimed.
	*	1928	The same	Entire population	For New York bank.
R. R. Doane	1935	1932	Capitalize income and tax payments	Income receipts	Apologist for present wealth distribution. Analysis deceptive, confusing.

* Published by W. Treseckow in 1931.

II American Studies of the Distribution of Income

1 HISTORY AND METHODS

Analysis of the distribution of income seems to have been secondary to, and certainly came later than, study of the distribution of wealth. Yet in discussion of social problems the former soon gained a significance not accorded the latter, and recent attempts to construct adequate distributions of income have been not only more numerous but also on the whole more successful than similar endeavors in the field of wealth distribution.

The problems encountered in constructing a distribution of income are similar to those faced in building a distribution of wealth. As before, there are two substantive elements: income and the receiving unit. The former has no single simple meaning. The money value of the total flow of economic goods emanating from wealth (both artificial and human) during a period such as a year is commonly referred to as national income, concerning which there is already a considerable body of academic literature, an imposing array of estimates, and an extensive

amount of government as well as public press discussion. But the total that is employed in constructing a distribution of income by size need not be, and for many problems should not be, the same as the total that is relevant as a comprehensive measure of the end-product of the economic system. The second element—income recipient—admits of as many definitions as the wealth-holding unit previously described.

As with wealth holdings, there has been no complete census of individual or family incomes in the United States. Therefore the problem is again one of raising a sample to universal coverage. It is chiefly in the nature of the samples, in the assumptions used in inflating them, and in the choice of income recipient that the various distributions of income differ.

a) *Spahr first to try*

As a sequel to his construction of the distribution of wealth, Spahr in 1896 essayed a distribution of income among families in the United States. There were four steps in his analysis.

Total national income was computed on the basis of Census returns and labor bureau reports of state and federal governments. Agricultural income was assumed equal to the 1889 value of farm product plus an estimate of the rental value of farm houses. Manufacturing income was derived from Massachusetts data on wages and profits, and from railroads and mines data in the Census reports. Service income was based on wages and profits in stores, while professional income was estimated from that of ministers and doctors. In estimating manufacturing and service income from wage rates (not earnings), average unemployment was allowed for in the following proportions: a dollar a day implied $260 per year, while $8 per week meant $360 per year. Income from urban real estate was estimated at 6 2/3 per cent of its value. The total income, prior to taxation, of the 22,735,000 persons gainfully employed in 1890 was finally set at $10,800,000,000.[45] This total was distributed among the 12,500,000 families in the United States.

[45] Spahr, *op. cit.*, Ch. V and VI, especially pp. 104, 105. This is more than a billion less than King's 1915 estimate of the national income in 1890 (see *Wealth and Income*, p. 132).

After enunciating the generalization that capital received two-fifths of the national income, and labor of all kinds the other three-fifths, Spahr declared that the 'safest guides' in the distribution of income by classes were the previously ascertained distribution of property, 'common observation' respecting the professional and business incomes of the wealthy and well-to-do, and Boston data on the distribution of rents.[46] On the basis of these guides, Spahr decided that the '$50,000 and over' class of wealth holders corresponded to the '$5,000 and over' class of income recipients, with the modification that 75,000 of the well-to-do families with possessions less than $50,000 were also in the '$5,000 and over' income category, thereby swelling the families in this group to 200,000. Similarly, the '$50,000–$5,000' class of wealth holders was assumed commensurate with the '$5,000–$1,200' class of income recipients. The above adjustment whereby 75,000 of the families in this well-to-do class were promoted to the '$5,000 and over' category left 1,300,000 families in the well-to-do group. Thus far the wealth class intervals have been converted to income class intervals, and the number of families in each class redistributed.

The transfer of 75,000 families to the wealthy class was accompanied by an increase of $2.5 billion in the wealth holdings of that group, while $1 billion (representing household goods) were subtracted from this figure, leaving a total of $34.5 billion. The wealth holdings of the well-to-do were reduced $4 billion because of this family shift and on account of household goods, while the wealth of the poorer class was cut $1.5 billion by the deduction of household goods. These classes therefore had left $19 billion and $7.5 billion, respectively. The return on this capital was estimated at 7 per cent for the wealthy and well-to-do classes, and 8 per cent for the poorer classes.[47]

The final step was to estimate the average labor income of the families in each class. That for the wealthy classes was set at

[46] Spahr, *op. cit.*, pp. 119–21. The reader is referred to Spahr's distribution of wealth presented in Sec. I. 1, b above, which is used as the starting point for the construction of his income distribution.

[47] *Ibid.*, pp. 125–8. No reason is given for the choice of these interest rates, and one wonders why the poor should enjoy a higher percentage return than the rich. Surely the poor are not better able to make wise and lucrative investments.

$3,500, for the well-to-do $1,200, and for the poorer $380, the last-named figure being a weighted average of an urban income of $500 and a rural income of $300. Although no precise method for estimating these averages is given, they are probably based on common observation, and made to jibe with the aforementioned dictum that labor of all kinds received three-fifths of the national income.

C. B. Spahr

DISTRIBUTION OF INCOME IN THE UNITED STATES, 1890 [48]

(PRIOR TO TAXATION)

FAMILY INCOME	FAMILIES *(thousands)*	AVERAGE INCOME, LABOR	CAPITAL *(millions)*	PERCENTAGE RETURN ON CAPITAL	TOTAL INCOME *(millions)*
$5,000 and over	200	$3,500	$34,500	7	$3,110
5,000–1,200	1,300	1,200	19,000	7	2,890
Under $1,200	11,000	380	7,500	8	4,800

It is possible to expand the resulting distribution, summarized in the accompanying table, by an added comment of Spahr's: "More than five-sixths of the income of the wealthiest class is received by the 125,000 richest families, while less than one-half of the income of the working-classes is received by the poorest 6,500,000 families." This statement has been introduced by the present writer into the foregoing table, and average family incomes computed, with the accompanying approximate results.

C. B Spahr

EXPANDED DISTRIBUTION OF INCOME IN THE UNITED STATES, 1890

INCOME CLASS	FAMILIES *(thousands)*	AVERAGE INCOME	TOTAL INCOME *(millions)*
$5,000 and over	125	$20,733	$2,592
	75	6,911	518
5,000 to 1,200	1,300	2,223	2,890
Under $1,200	4,500	556	2,500
	6,500	354	2,300 [49]

A. J. Ferris, a Philadelphia writer with pronounced preconceptions, cast Spahr's income distribution, by a series of unexplained adjustments and assumptions, into a different form.

[48] *Ibid.*, p. 128.

[49] This allocation of "less than one-half" is arbitrary.

A. J. Ferris

APPROXIMATE DISTRIBUTION OF INCOME IN THE UNITED STATES[50]
(PRESENT STATUS)

(INDIVIDUAL YEARLY INCOMES)

INCOME CLASS	PERSONS (*thousands*)	INCOME (*millions*)	AVERAGE INCOME
Under $60	15,000	$ 525	$ 35
60– 125	35,000	2,975	85
125– 250	10,000	1,750	175
250– 750	3,500	1,625	450
750– 2,500	1,200	1,500	1,250
2,500–10,000	250	1,250	5,000
$10,000 and over	50	775	15,500
All Income Classes	65,000	$10,400	$ 160

Ferris, in a note to this table, states: "The present classification into several divisions is an amplification of Dr. Spahr's, following the data given in his book when they cast any light on the subject, and for the rest simply based on probability and the analogy of the main classification. The results here given have been submitted to Dr. Spahr, and in their general features were approved by him." Because of the transition from a family to a person basis, it is hard to draw conclusions concerning differences in the shapes of Spahr's and Ferris' distributions. And since this transition—ordinarily a difficult and treacherous statistical job—is not explained, no judgment concerning its validity is possible. Confidence in Ferris' adjustments, however, is not encouraged by the nature of his proposal for alleviating the existing inequality in the distribution of incomes. Ferris would increase everyone's income by $160, the approximate amount of the average income in 1890. Such a step, he reasons, would double prices, make each man's real income equal one-half of his former monetary income plus one-half of the $160 average income, and thereby reduce those incomes above the average and increase those below the average. Such a naïve suggestion is typical of Ferris' book.

b) *Streightoff shied away*

A fine sense of caution, statistically, characterized the next student of the problem, F. H. Streightoff. In his 1912 study of the

[50] A. J. Ferris, *Pauperizing the Rich* (Philadelphia, 1899), tabulation facing p. 192. 'Present Status' presumably refers to 1890, or shortly thereafter.

Distribution of Incomes in the United States,[51] he was concerned primarily with pointing out the utility of income statistics, the available American data on incomes, and their insufficiency for the construction of a complete distribution of income. He tried to derive the distribution of income from property, but finally concluded: (a) the number of persons receiving income from capital is large but unknown, (b) the total national income from capital cannot be accurately determined, (c) the distribution of income from property is a futile quest. Although he realized that value of farm product was not equal to net farm income, and that wage rates were not distributed in the same fashion as earnings, Streightoff did employ such figures to construct a "distribution of incomes primarily from labor". His principal sources were the *Eighteenth Annual Report* of the Commissioner of Labor (1903), the *Censuses of Mines and Quarries* (1902), *of Manufactures* (1905), and *of Agriculture* (1900), Kansas *Bureau of Labor Reports* (1903–07), and the *Annual Minutes* of ten typical Methodist Episcopal Church conferences (1910). His resulting table distributed among three income classes the 19,658,000 males 16 years and older gainfully occupied in the United States in 1904, including industrial workers, ministers, agricultural laborers, and heads of farm families. Since neither the income received by each class nor by the total group was estimated, a frequency distribution in Lorenz curve form of Streightoff's results is not practicable.

c) *King again the pioneer*

As in the field of wealth distribution, so in that of income distribution, W. I. King did the pioneer work as far as statistical adequacy is concerned. In his 1915 study he agreed with Streightoff "that it is, at present, impossible to give any accurate picture of the distribution of incomes among the population as a whole." [52] However, he had some Wisconsin income tax data not available to Streightoff, so he attempted "to classify roughly the twenty-eight millions of families living in the Continental United States according to the income which each, respectively,

[51] *Columbia University Studies*, Vol. III, No. 2 (New York, 1912), especially pp. 46–56, 137, 150.
[52] *Wealth and Income*, p. 219.

receives." [53] King took as granted that "any classification of income must, necessarily, be based upon receipt of families rather than individuals for it is by families that incomes are received and disbursed".[54] Although the methods followed by King in constructing his distribution of income among families in 1910 "were mainly graphic and were too varied to describe here",[55] they may be grouped into three divisions.

Wisconsin income tax data compiled by H. M. Trumbower were used to solve the question of how middle class incomes were distributed. Wisconsin was considered a "peculiarly good sample state" with a per capita wealth "about equal to the average for the United States as a whole". Therefore, the central part of the curve for Wisconsin was considered "fairly representative for the middle class throughout the entire nation".[56]

The incomes of the wealthy were inferred from United States Treasury Department and Congressional estimates of the incomes of the very rich in certain metropolitan centers in the East, and from preliminary reports on the federal income tax.

Lower class incomes were estimated on the basis of Census data, reports of the United States Commissioner of Labor, and investigations by the bureaus of labor of the various states. These were supplemented by private studies of workingmen's budgets.

The results of these three methods were combined in an unexplained fashion to give "The Estimated Distribution of Income among the Families of the Continental United States in 1910".[57] The fifty class intervals included family incomes between $0 and $50,000,000. For incomes under $1,400 the recipients were classified as 'single men', 'single women', and 'men or widows with families'; for incomes over $1,400 the only unit of income recipient was the family.

No explanation of how King estimated the $30,529,000,000 of national income distributed in the foregoing fashion among the 28,000,000 families is attempted here, primarily because, with King and his successors, the problem of estimating the total na-

[53] *Ibid.*, p. 217.
[54] *Ibid.*, p. 222.
[55] *Ibid.*, p. 221.
[56] *Ibid.*, p. 220.
[57] *Ibid.*, pp. 224–6.

tional product was of prime, not secondary importance.[58] Spahr and Streightoff considered the derivation of the national income as a means to an end, the goal being the family or individual distribution of this product. Since Streightoff, an accurate formulation of the value of our national output of economic goods has been emphasized as an end in itself. The development of the latter technique is outside the scope of this paper, so in discussing King and subsequent writers, the national income total to be divided among the individual claimants will be considered given data.

d) *Macaulay's distribution for 1918*

The first publication of the National Bureau of Economic Research was a two volume study, *Income in the United States*.[59] As a collaborator in this work F. R. Macaulay made a thorough analysis of the frequency distribution of annual income among personal income recipients in the United States in 1918.[60] Income was defined to be money income plus those items of commodity income on which a money value is placed, such as rental value of owned houses and value of farm produce consumed by farmers' families.[61] Income recipient was taken to be the individual and not the family because (a) it is the individual who comes into direct economic relationship with the machinery of distribution, and (b) use of families still leaves unsolved the question whether to employ theoretical families, biological families, or families expressed in a need-unit such as the 'ammain'.[62]

[58] King does not state specifically whether his concept of income includes capital gains and losses. From his methods and sources, however, one gathers that it does not.

[59] W. C. Mitchell, W. I. King, F. R. Macaulay, and O. W. Knauth, Vol. I (1921), summary, and Vol. II (1922), details.

[60] *Ibid.*, II, 341–425.

[61] Income is defined to include also statutory capital gains and losses, since these were apparently not extracted from the income tax data before building up the tail of the distribution.

[62] *Ibid.*, II, 341, 342. Macaulay does not deny that the family is the chief unit of economic need. He apparently takes the term 'ammain' from an article by Edgar Sydenstricker and W. I. King, 'The Measurement of the Relative Economic Status of Families', *Journal of the American Statistical Association*, XVII (1921), 842–57.

The total income estimated by the National Bureau was distributed, before deduction for taxes, among all who had money income as follows: [63]

1. Income tax data, unusually complete for 1918, were adjusted to include (a) farmers and small business men who filed no returns, (b) evasion by reporting persons, (c) non-monetary income referred to above, (d) income from tax-exempt securities.

2. O. W. Knauth's distribution of incomes above and below $2,000, another part of this National Bureau study, was used as a check on Macaulay's distribution.

3. Incomes under $2,000, inadequately covered by income tax statistics, were estimated in an unexplained fashion on the basis of wage distributions, small samples of farmers' incomes, and other studies such as A. T. Emery's unpublished sample of Chicago incomes.

4. Since some business men incur net losses, Macaulay estimated the number and amount of these negative incomes, and spread them in some manner throughout the distribution.

5. The final frequency curve was smoothed on the assumption that, even though the distribution of income followed no mathematical law,[64] nevertheless it would not be bimodal and 'bumpy'.

The final distribution was presented with small class intervals ranging from 'under zero' to '$4,000,000 and over'. The 2,500,000 soldiers, sailors, and marines in 1918 were excluded from the number of income recipients on the assumption that in peacetime their incomes would be distributed similarly to those of the others.

c) *King a prolific contributor*

Before the publication of the next major work in this field by the Brookings Institution in 1934, a dozen years passed marked by

[63] Mitchell, King, Macaulay, and Knauth, *op. cit.*, I, 124–6. "Personal income recipient" here corresponds "closely to the Census expression person gainfully employed . . . . Perhaps the most important difference is that we do not, and the Census does include as separate income recipients, farm laborers working on the home farm." (*Ibid.*, II, 342n.) But what about those persons who, although not 'gainfully employed', are nevertheless in receipt of income (e.g., from property)?

[64] His examination of Pareto's famous law led him to this conclusion, *ibid.*, II, 393, 394.

increasing interest in the interpretation of income distributions, especially by consumption economists and marketing students.[65] Meanwhile King worked out two complete distributions of income among individuals, one for 1921 and another for 1928, and attempted to trace annual changes in the distribution of income for 1914–26.

A distribution of personal incomes in 1921, comprising 32 class intervals from $0 to $1,000,000, was made at the National Bureau of Economic Research by King, left unpublished, and in 1934 utilized in percentage form by Maurice Leven in *America's Capacity to Consume*.[66] No details whatsoever were given by Leven as to how this distribution had been constructed, but the circumstances surrounding the 1921 distribution have already been set forth (above, footnote 18), and examination of the unpublished manuscript describing its construction reveals that the work fell into four stages. First King derived the distribution of earnings among employees; then he distributed the income of farmers; the third step was to find the income distribution of non-farm entrepreneurs and income recipients not gainfully employed; and the final stage was to combine these distributions into one of income among all classes. Each step will be described in turn.

Employees were construed to be not only wage earners but also salaried workers, including highly paid executives. Of the $34.3 billion 1921 wage and salary bill, $22.7 billion went to the former (23,602,469 persons) and $11.6 billion to the latter (7,137,531 persons). This was allocated by means of sample wage distributions for earnings under $2,000 and *Statistics of Income* data for earnings over $2,000. The final earnings distribution was a composite of 132 sample distributions (weighted according to importance and adjusted to 1921 conditions) for the lower classes

[65] E.g., Hazel Kyrk, *A Theory of Consumption* (Boston, 1923), E. E. Hoyt, *The Consumption of Wealth* (New York, 1928), W. C. Waite, *Economics of Consumption* (New York, 1928) and P. H. Nystrom, *Economic Principles of Consumption* (New York, 1929).

[66] Maurice Leven, H. G. Moulton, and Clark Warburton (Brookings Institution, 1934), pp. 177, 182–4. Leven derived 1929 equivalents from the 1921 figures in order to provide a check on his own computation of a distribution of income for 1929. It could constitute such a check, it should be noted, only to the extent that the inequality in the income distribution had not changed from 1921 to 1929.

and of *Statistics of Income* frequencies for the higher ranges. After constructing this distribution, King proceeded to break it down by sex and industry. He also ventured an *obiter dictum:* inequality is not due solely to income from property: earnings themselves are decidedly unequal, not only at the extremes but all along the earnings scale.

The Bureau of Agricultural Economics constructed a distribution of income among a sample of farm crop reporters in 1922 which formed the basis of King's distribution of income among farmers in 1921. This sample curve was adjusted so that an income total computed on the basis of its shape would correspond with the farm income totals derived by the National Bureau of Economic Research. Current money income was first distributed according to the crop reporter sample. Current money income was then supplemented by imputed interest on consumption goods owned, to give entire or total current income. Current money income was also supplemented by the value of commodities produced and consumed on the farm, to give current money and commodity income. Finally, this was corrected for changes in the value of property owned (i.e., unrealized capital gains and losses), to give total money and commodity income. This last adjustment was of no mean proportions: an entire current money and commodity income of $4.4 billion was slashed to a total money and commodity income of $2.4 billion. This $2 billion decline represents the diminution in the command over consumption goods of the sum of money representing the value of farm property. King argued that this was a real not a nominal loss. This total money and commodity income was distributed in two distinct fashions, and the resulting distributions combined by simple averaging. In the first, the 1922 crop reporter curve was adjusted to fit the revised income total. In the second, the distribution of current money income, itself based on the crop reporter sample, was used as a datum from which was subtracted (or added) the total losses (or gains) of farmers arising from changes in the value of their farms, livestock, machinery, etc., as given by the Census. This process assumed that those farmers possessing the most property suffered the heaviest property losses when farm prices fell. The average of these two distributions of total money and commodity income was supplemented by im-

puted interest on consumption goods owned to form the final distribution of total income of farmers (including commodity income and imputed interest).

Having estimated the earnings of employees and of farmers, King proceeded to distribute the remaining income among those persons who were neither employees nor farmers, i.e., the non-farm entrepreneurial group. The incomes of those in this group receiving less than $2,000 were assumed to be distributed similarly to the earnings of employees. The incomes of those above the $2,000 scale were distributed by *Statistics of Income* tables. Once the current money income was distributed, King next corrected for changes (measured by consumption goods prices) in the value of property held by persons in each income class, in order to get total income.[67] The distribution of dividends, interest, and rents, as revealed by *Statistics of Income,* was used as a basis for apportioning gains in the value of property. The final adjustment of this distribution among non-farmers was for negative incomes. This was done by "the device of drawing a smooth histogram through the records of persons having positive incomes and extending the same free-hand into the negative side of the field". The statistical manipulations involved in this third stage of the income distribution analysis were exceedingly complicated and oftentimes quite arbitrary.

The synthesis of King's analysis appears in the last chapter of the section dealing with income. Here he combined the three distributions arrived at in the preceding steps, and got a series of distributions for all the income recipients in the United States, based successively on (a) current money income, (b) current money and commodity income, (c) total current income (including imputed income from consumption goods owned), (d) total income (including gains or losses in the value of property owned, i.e., unrealized capital gains and losses). These distributions, in the order named, acquired an increasingly comprehensive concept of income. The final one, i.e., of total income (d), is preferred, said King, if one views the subject of income distribution from a "technical financial standpoint". The monograph con-

[67] Thus total income includes not only realized capital gains and losses (already comprehended by current money income), but also unrealized capital gains and losses.

cluded with a Lorenz curve comparison of King's total income and current income distributions with Macaulay's 1918 distribution. The inequality thus indicated decreased in passing from one to the next of these distributions in the order listed.

The second of these distributions of King's, among individuals in 1928, has had a curious history. Never published by King, it was taken over by Leven, converted to a family basis, and incorporated in *The Ability to Pay for Medical Care*.[68] It was then seized upon by Louis Bader who condensed it from twelve to five class intervals, computed the percentages of families and incomes in each category, applied these percentages to the total number of families and amount of income in 1932, and then analyzed what happened to family expenditures from prosperity to depression.[69] This work posited that the 1932 national income was distributed in dollars in the same way as in 1928, which Bader claimed to be "a fair assumption since all income groups have suffered, due to decreases in all forms of income".[70] Although this assumption may be legitimate for Bader's purposes, it begs the question generally asked—Does inequality of income distribution change from prosperity to depression?—so this survey will ignore the 1932 distribution, and consider only the source from which it was derived.

King's estimate of the distribution of individual income recipients according to amount of annual income in 1928 was originally constructed for the Central Hanover Bank and Trust Company of New York. No details concerning the statistical devices utilized in its synthesis have been published, although King, in a personal letter to the present writer, states it was made along lines similar to his 1921 estimate except that "the figures for the lower income classes are . . . merely rough approximations" since his sponsor "was not interested in the distribution in the lower brackets".[71] Leven, then on the staff of the Committee on the Costs of Medical Care, used it to derive an estimated distribution of families according to annual income in 1928. The work

[68] L. S. Reed, *The Ability to Pay for Medical Care* (Committee on the Costs of Medical Care, University of Chicago Press, 1933), pp. 10, 11.
[69] 'The American Family Income and Prosperity', *Journal of the American Statistical Association*, XXVIII (1933), 303–11.
[70] *Ibid.*, p. 305.
[71] Letter dated April 4, 1938.

done by Leven in making this conversion seems to have been a testing ground for the procedure he later employed in the Brookings study. On the basis of the then incomplete 1930 Census returns, Leven estimated [72] the number of families with one gainful worker, with two, three, and four gainful workers. These gainful workers were broadly classified into main breadwinners and supplementary earners, the latter group composed largely of gainful workers under the age of twenty. The problem then became one of breaking up King's distribution of individual income recipients and recombining the component parts into family units. The first step in this process was to divide the individual income recipients into supplementary earners and chief breadwinners. This was done by assuming that (a) all persons under twenty were additional earners in families headed by others, (b) their incomes were all under $1,200, (c) most of the female workers were supplementary income recipients, and (d) their incomes were distributed according to certain data collected by Leven in a New York City survey.[73] The residual distribution resulting from subtracting minors and females from King's distribution was taken to represent chief breadwinners and persons living independently outside family units. The incomes of this second group were assumed to be distributed in the same manner as those of main breadwinners. Leven's second step was to allocate the supplementary earners to the families having such members. The procedure was expressed thus: "the income of each head of family in a given income class was combined with the income of a supplementary earner picked in accordance with the probability represented by the supplementary earner's income curve". This required the assumption "that the probability of a main breadwinner being associated with a supplementary earner of a given income was the same for all incomes of the main breadwinners, and that this probability is represented by the income distribution of the supplementary earners".

Although the use of the word 'probability' here is a little confusing if one tries to attach a technical mathematical meaning to

[72] 'Note on the Distribution of Income', Appendix A of Reed, *op. cit.*, pp. 99–101.
[73] Maurice Leven, *The Incomes of Physicians* (Committee on the Costs of Medical Care, University of Chicago Press, 1932), p. 127.

it, a conceivable interpretation of Leven's description would imply the following procedure. Assume income distributions:

OF BREADWINNERS IN FAMILIES HAVING ONE SUPPLEMENTARY EARNER		OF SUPPLEMENTARY EARNERS		
INCOME CLASS	NUMBER	INCOME CLASS	NUMBER	PERCENTAGE
0–$500	100	0–$500	300	50
500–1,000	150	500–1,000	180	30
1,000–1,500	200	1,000–1,500	120	20
1,500–2,000	150			

Under this interpretation, Leven allotted (a) supplementary earners in the 0–$500 class to 50 per cent of the breadwinners in each income class, (b) supplementary earners in the $500–$1,000 class to 30 per cent of the breadwinners in each income class, and (c) supplementary earners in the $1,000–$1,500 class to the remaining 20 per cent of the breadwinners in each income class.[74]

The final distribution obtained in this manner is the sum of the income frequencies of the several groups of families. It also includes families without gainful workers, allocated a $1,200 income.

Abandoning for the moment his elaborate statistical devices for deriving a complete distribution of income, King in 1930 used *Statistics of Income* data and the Census figures for gainfully employed to derive truncated distributions of income for 1914–26.[75] The distribution of income recipients above $5,000 was taken directly from *Statistics of Income* for each year, while the rest of the gainfully employed were put in the 'under $5,000' category. All distributions were reduced to 1913 dollars by means of indices of the average prices of consumption goods used by different income classes of the population. Such an analysis of

[74] This illustration requires explanation in at least two points. First, supplementary earners are 'allotted' to the main breadwinners by shifting the income curve of the main breadwinners to the right by an amount equal to the average income of the supplementary earners in each income class. Second, the present illustration relates to merely one set of supplementary earners, e.g., those in families with only one supplementary earner. A similar process would have to be gone through again in the case of families with two supplementary earners, three, etc. This explains why the number of supplementary earners was taken equal to the number of breadwinners in the illustration.

[75] *The National Income and Its Purchasing Power* (National Bureau of Economic Research, New York, 1930), Ch. VII.

Statistics of Income data is really not a construction of a complete distribution of income by size, as we have been using the term, for the 'under $5,000' class comprising an overwhelming majority (about 97 per cent) of the income recipients was not subdivided. Furthermore, no account seems to have been taken of those without gainful employment who were nevertheless in receipt of income. Other students have made similar partial analyses. For example, W. L. Crum applied Pareto's graphic method to *Statistics of Income* data without any attempt to construct the distribution of income below the income tax exemption point.[76] The Pareto slopes he computed, therefore, applied only to the tail of the distributions. N. O. Johnson, in a defense of Pareto's thesis, made a similar study of inequality in the upper brackets.[77] And M. A. Copeland analyzed, on the basis of federal income tax data, the problem of inequality from a different angle, namely, per capita income, and per cent of total income received by the wealthiest 10 per cent of income recipients.[78]

f) *Leven's distribution*

A widely publicized attempt to construct a distribution of income by size for the United States was made by the Brookings Institution in 1934 as an integral part of the second volume in its study of the distribution of wealth and income in relation to economic progress. Leven, in charge of this part of the study, constructed a comprehensive distribution of income among families for 1929, which comprised twenty-seven class intervals from "under $0' to '$500,000 and over'. The method followed was long and involved, and only its outline can be sketched here.[79] Leven first converted Macaulay's estimate for 1918 and King's unpublished figures for 1921 into 1929 equivalents, and then used these two distributions as checks upon his own independent construction of the distribution of income among individuals in

[76] 'Individual Shares in the National Income', *Review of Economic Statistics*, XVII (1935), 116–30.

[77] 'The Pareto Law', *Review of Economic Statistics*, XIX (1937), 20–6.

[78] 'The National Income and Its Distribution', *Recent Economic Changes* (National Bureau of Economic Research, 1929), II, 757–839; see especially pp. 833–7.

[79] The final distribution is presented in Leven, Moulton, and Warburton, *op. cit.*, p. 54. The calculations are presented and methods explained in Ap. A, 'Income and Its Distribution', pp. 137–238.

1929.[80] Leven's independent estimate was arrived at roughly as follows:

Earnings were distributed among gainfully employed non-farmers on the basis of federal income tax statistics and various sample distributions weighted by their importance and adapted to 1929 conditions. This distribution was adjusted so that aggregate earned income equaled the Department of Commerce estimate of total occupational income for 1929. It was then converted into one of total income by the use of previously ascertained ratios of total income to occupational income. Finally, to this distribution was added the estimated distribution of income recipients without a gainful occupation. In making this union it was assumed that "the distribution of those without gainful occupations was like that of the individuals with gainful occupations".[81]

For farmers, the first task was to estimate total income, and then distribute this total. Net farm income was derived from Department of Agriculture figures, and the distribution was made on the basis of (a) Census figures of 'Value of Farm Products' for individual farms, and (b) samples that showed the relation between gross and net income of individual farmers.

The distributions for non-farmers and farmers were apparently added to give the final distribution of personal incomes in 1929. The next step was the conversion of this distribution among persons into one among families. In an unexplained fashion the personal distributions of farm and non-farm incomes (treated separately) were each broken down into a threefold frequency distribution of personal incomes (a) for all heads of families of two or more persons, (b) for supplementary income recipients, (c) for unattached individuals. Parts (a) and (b) were then combined to make a distribution of family incomes. The distribution of families with only one income recipient followed readily from the assumption that its form was the same as that of heads of families having any specified number of supplementary earners (i.e., each frequency in part (a) was multiplied by the ratio of the total number of families with one income recipient

[80] *Ibid.*, pp. 177–84. It has already been noted that the utility of such a check varies directly with the stability of income inequality between those years.
[81] *Ibid.*, pp. 185, 186.

to the total number of all families). "The residual frequencies, obtained by subtracting the distribution of one income families, constituted the distribution of principal incomes in families of two or more income recipients." [82] To this Leven added an equal number of supplementary incomes (part (b) above) to obtain the combined distribution of the first two income recipients. This was divided in an unexplained way into families having two income recipients and families having more than two. The latter distribution was adjusted to include a third income recipient for each family, and the process was repeated until distributions for all five groups were set up.

All this may seem complex, but the complications are not yet at an end. Families with more than one income recipient were distributed over the income classes in the same proportions in which the supplementary incomes were distributed; then the distribution curve of principal incomes was shifted to the right along the income scale by amounts equal, for each class interval, to the corresponding class average of the supplementary incomes. The several distributions thus obtained were plotted as cumulative curves and then added to give a composite distribution incorporating families with one and two income recipients. A similar process was employed in combining the third, fourth, and fifth income recipients with the basic distribution. All this mathematical juggling was used only for incomes under $15,000; families with incomes over $15,000 were assumed to be distributed proportionately to principal incomes.[83]

In the end we have a distribution of income by theoretical families of two or more persons, with capital gains and losses included in the concept of income, and with the twenty-seven class intervals ranging from 'under $0' to '$500,000 and over'.

g) *Tucker on inequality*

Two major contributions have been made very recently to the study of the distribution of income by size—one by R. S. Tucker, another by the National Resources Committee. The first ap-

[82] *Ibid.*, p. 224.

[83] *Ibid.*, p. 226. The assumptions implicit in Leven's analysis are effectively singled out by A. F. Burns, in 'The Brookings Inquiry', *Quarterly Journal of Economics*, L (1936), 495, 496.

peared in the August 1938 issue of the *Quarterly Journal of Economics*.[84] It is limited to income tax data, and therefore would have been accorded, in this report, space similar to that given the Pareto analyses of King, Johnson, and Crum were Tucker's article not distinguished from these predecessors in several respects. In the first place, Tucker attempts to carry the picture back to the Civil War. Second, his analysis of the existing income tax data is relatively intensive, several measures of inequality other than the Pareto slopes being used. And finally, he boldly asserts as an introductory thesis that changes in the income distribution of the well-to-do indicate what is probably happening to the rest of the distribution, since the two ratios of (a) income of the wealthy to income of the middle class, and (b) income of taxpayers to income of non-taxpayers, are approximately identical.

Tucker differentiates three concepts of income: legal income, which conforms to the statutory definition of income, with such adjustments as are necessary to maintain comparability; spending power, which equals legal income plus tax exempt interest minus the income tax paid; earning power, which equals legal income minus realized capital gains plus realized capital losses plus tax-exempt interest plus gifts, charitable contributions, and the like.

Statistics of Income data for 1914–36 are analyzed by means of five measures of dispersion. The first two are Pareto slopes, one referring to the number of persons and the other to amount of income, each being taken above the $5,000 income level and cumulated by income class from top to bottom. The third measure is the arithmetic average of all incomes above $5,000. If $5,000 is taken as the modal income, comparison of this average with $5,000 suggests the skewness in the distribution. The fourth indicator is one apparently introduced by Hans Staehle [85] and is the ratio of the cumulative median income (the income such that

[84] 'The Distribution of Income among Income Taxpayers in the United States, 1863–1935', *Quarterly Journal of Economics*, LII (1938), 547–87.

[85] 'Short-Period Variations in the Distribution of Income', *Review of Economic Statistics*, XIX (1937), 133–43. (Cited by Tucker.) This measure was foreshadowed by Holmes ('Measures of Distribution', *Publications of the American Statistical Association*, III (1892–93), 141–57) who suggested using the difference between the median lines for wealth owned and for number of wealth owners.

individuals with greater incomes receive fifty per cent of the total income) minus the median income to the cumulative median for all incomes above $5,000. Its lower limit, 0 per cent, is absolute equality, and its upper limit, 100 per cent, implies that all the income is received by the upper half of the income recipients in the group. The picture painted by these four measures is checked by a composite indicator comprising the 'earning power' income received by all taxpayers (above the $5,000 level) minus income taxes paid, divided by national income paid out. The results of applying these various measures to the different concepts of income led Tucker to the belief that there has been an increased diffusion of income over the twenty-three years studied. Therefore he next addressed himself to the question, how long has this increasing diffusion been going on?

The income tax law of 1894 yielded scanty data with which to essay an answer to this question, because it was declared unconstitutional before it became fully effective. The Civil War income tax laws, however, yielded official published statistics which when supplemented by various private lists [86] made it possible for Tucker to employ two of the aforementioned measures of concentration: the first, referring Pareto slopes to number of recipients, and the third, being the arithmetic average of incomes above $2,000 and then above $3,000. The results of this analysis, together with the fact that reportable income in the 1860's did not include interest and dividends from public companies and from government bonds, or certain realized capital gains (items which normally accrue to the wealthy and whose exclusion would therefore understate the concentration of in-

[86] Tucker cites J. A. Hill, 'The Civil War Income Tax', *Quarterly Journal of Economics*, VIII (1894), 416–52, 491–8, for general information on these Civil War data. The two private lists cited are *Income Record* (New York, 1865) and *Income Tax of Residents of Philadelphia* (Philadelphia, 1867). Another that he failed to cite is *Income Tax of Residents of Philadelphia and Bucks County* (Philadelphia, 1865). All these tax lists are anonymous. The first gives the taxable income for 1863 of every resident of New York. Unlike the other lists, this one contains a 'Publisher's Preface' which discusses such topics as the practical significance of a distribution of incomes, the English income tax, and tax evasion (estimated at not more than 10 per cent). The second list describes 'The Rich Men of Philadelphia' by size of income in 1865 and in 1866, and is based on the latest returns filed by August 1867. The third list classifies the same personages by size of income for the year ending April 30, 1865.

come), led Tucker to the belief that incomes were less concentrated since 1916 than in Civil War days. Tucker concludes his analysis with a brief survey of the shifting composition of the wealthy group. He reviews the results of three studies that have been made of this problem:

1. 'Investigation of Bureau of Internal Revenue', *Senate Report* no. 27, 69th Congress, 1st Session, Part 2. This traces 6,633 individuals with incomes over $100,000 in 1916.

2. Edward White, 'Income Fluctuation of a Selected Group of Personal Returns', *Journal of the American Statistical Association*, XVIII (1922), 67–81. The 1,636 individuals or estates with incomes over $300,000 in any of the years 1914–19 are traced.

3. Bureau of Internal Revenue, *Statistics of Income*, 1922 (Washington, 1925), pp. 11–15. This follows the fortunes of 1,296 individuals with incomes over $300,000 in any of the years 1916–22.

These studies all indicate, Tucker avers, that persons in "the upper income classes have been a very shifting group",[87] although some of this shifting after 1916 may have been due to sharing of taxpayer's property with wives and children in an effort to qualify in the lower tax brackets.

Unlike most of the other income distribution students whose works we have examined in this paper, Tucker winds up his contribution with several general conclusions. In the first place, fluctuation in the concentration of wealth, during the business cycle, is less than in the concentration of income. Second, the concentration of income increases during prosperity and decreases during depression. Third, the size of the national income is the important consideration. Fourth, bank reforms are needed "to prevent excessive use of credit". And finally, the question is not how large are incomes, but whether they are the result of activities beneficial or harmful to the nation.[88]

[87] Tucker, *op. cit.*, p. 583. It is significant to note, however, that he seems to take no account of deaths among these taxpayers. Surely a fair number left the picture for that reason.

[88] *Ibid.*, pp. 585–7.

h) *Family incomes in 1935–1936*

The combination of a large new sample of family incomes and the interest of a government agency in the income distribution problem have resulted in the most reliable as well as the most recent distribution of income—that presented in the report of the National Resources Committee, entitled *Consumer Incomes in the United States, Their Distribution in 1935–36*, prepared by Hildegarde Kneeland and her staff, and dated May 27, 1938. The Study of Consumer Purchases, a Works Progress Administration project conducted by the Bureau of Home Economics and the Bureau of Labor Statistics, with the cooperation of the National Resources Committee and the Central Statistical Board, was a nationwide canvass of 300,000 families that provided not only the initial impetus but also the basic material for this latest distribution of income.

The National Resources Committee report may be divided into three sections: a 36-page summary, a detailed appendix on 'Sources and Methods Used in the Study', and a concluding compendium of 'Statistical Tables for Reference Use'. Although the distribution is largely grounded on the Consumer Purchases data, other samples for single men and women, earnings figures and federal income tax statistics were used. Since the Consumer Purchases Study covered family incomes for a year ending between December 1935 and December 1936, the National Resources Committee distribution is taken to refer to the year ending June 30, 1936, the period covered by the majority of the schedules. Comprehended by the distribution are 29,000,000 families of two or more persons, 10,000,000 single individuals living alone or as lodgers. Classified separately are 2,000,000 persons living in institutional or semi-institutional groups.

As a first step, the Consumer Purchases data were divided into 729 homogeneous family groups—homogeneous in respect of size and occupation of family,[89] relief status,[90] color and nativity, size of community, and geographical region. All families in the

[89] A family's occupation was determined on the basis of the source from which the family received the largest amount of income.

[90] Relief families were segregated and considered separately, since they could not be classified into as many homogeneous groups as the other families.

United States were similarly split up by means of Census returns, and the percentage distribution (by size of income) of each segment in the Consumer Purchases sample was applied to the corresponding segment of the 29,000,000 families. These components were supplemented by means of federal income tax data for incomes over $7,500, after which the parts were summated to give the estimated national distribution of family incomes. 'Income' in this part of the study includes both money and non-money income, net after business expenses and business taxes, but before income, poll, and sales taxes. Federal income tax data, used to construct the tail of the component distribution, were first adjusted by removing capital gains, by adding interest paid, capital losses, taxes, contributions, and tax-exempt interest received, by combining separate returns of husbands and wives, and by making allowance for understatement and non-reporting of income.

The number of families in each income class above $7,500 was derived wholly from income tax data. This distribution was then tacked bodily onto the distribution based on Consumer Purchases data. Since the population weights used in constructing the latter distribution had accounted for all families in the United States, the addition of the income tax 'tail' resulted in an overstatement of the number of families. In a manner that left unchanged the shape of the distribution, this excess number of families was subtracted from the income classes under the $7,500 level.[91] The resulting distribution contains twenty-eight income classes ranging from 'under $250' to '$1,000,000 and over'.

The distribution for single individuals was built up by means of a more tenuous procedure and is therefore less reliable. The distribution for non-relief single women is based largely on data resulting from studies by the United States Women's Bureau and the United States Employment Service. The distribution of non-relief single men was derived from this distribution of non-

[91] *Consumer Incomes in the United States*, pp. 85, 86. For a given income class, the percentage that the number of families in that class bore to the total number of families with income less than $7,500 was applied to the number of 'extra' families, and the resulting product subtracted from the number of families in the given class. By this procedure the percentages of families in each class under $7,500 were left unchanged; the number of families in each class had been reduced proportionately to the frequency of that class.

relief single women by using the relationship known from various studies to exist between earnings of men and women. These two distributions were then checked by small samples from the Consumer Purchases Study and the National Health Survey. For single individuals who received relief at some time during the year, fragmentary Works Progress Administration data and certain assumed relationships between incomes of relief and non-relief individuals were employed. The distributions for relief and non-relief individuals were then combined to give a composite distribution for single persons. This series of frequencies was also supplemented, above the $3,000 income level, by federal income tax data. The income class intervals are identical with those in the family distribution.

Institutional residents presented difficulties not raised by either families or single individuals, since much of their income is in the form of food, clothing, and shelter provided through a central commissary. Civilian Conservation Corps incomes were distributed with the help of data supplied by the Director. Enrolees were credited only with that portion of their monetary income not sent home to their parents. Incomes of Army and Navy personnel were distributed by means of data embodied in pertinent Congressional Committee hearings. A combination of these two distributions—Civilian Conservation Corps and Army and Navy—was made and the resulting percentage frequencies applied to incomes of workers in labor camps and crews on vessels. For the other institutional residents, reports of various state welfare departments were used in devising the distribution. Such residents were assumed to have incomes equal to average subsistence costs, exclusive of administrative overhead and capital outlays. No composite distribution for institutional residents is presented because the institutional group rather than the constituent thereof makes up the spending unit. For the same reason the incomes of institutional residents are excluded from the composite distribution of incomes of all consumers (i.e., of families and of single individuals). The income distribution of families and single individuals combined contains twenty-eight class intervals ranging from 'under $250' to '$1,000,000 and over'. The resulting figure for aggregate income received is 5 per cent less than the Department of Commerce estimate for income

paid out, after appropriate adjustments for the sake of comparability. Considering the fundamentally different nature of the two independent estimates, one can agree with Miss Kneeland that "this discrepancy does not appear excessive".[92]

In addition to this overall distribution, the distribution of family incomes is further subdivided by size of family, region, size of community, occupation, and color. All in all, this monograph presents a relatively comprehensive picture of the distribution of income by size in the United States.

A partial distribution of income, the full details of which are lacking, has been presented by L. J. Chawner in a National Resources Committee monograph entitled *Residential Building*. It covers nonfarm households alone and applies to 1933 since its lower ranges are based on D. L. Wickens' *Financial Survey of Urban Housing*, a Department of Commerce publication. In common with the other distributions of income it, too, is based on *Statistics of Income* tabulations in the income classes above the $5,000 level.

The number of nonfarm households was first estimated from the 1930 Census and then extrapolated to 1933 conditions by means of the Census Table 1 on annual population increases, after allowing for the doubling up of families during the depression years. These households were then distributed by the income class frequencies indicated in the *Financial Survey* and in *Statistics of Income*. Because the resulting aggregate of nonfarm income was slightly less than the corresponding national income estimate of the Department of Commerce,[93] the distribution was adjusted upward until the two totals agreed.

2 PURPOSES OF STUDIES

The forces that motivated Spahr and King to write their books on income distribution have already been discussed in the sections on wealth distribution. It deserves to be repeated here that Spahr recognized the distribution of income as intrinsically more

[92] *Ibid.*, p. 35.

[93] The Department of Commerce figure was adjusted for agricultural income, for incomes of individuals living in boarding houses, hotels, and labor camps, for net capital losses, for dividends to insurance policyholders, etc.

important than the distribution of wealth, while King is credited with the statement that "income is the best single criterion of economic welfare".[94] Spahr's principal interest was in the social problems involved rather than in the refined statistical methods necessary, while King, though perhaps primarily interested in the statistical aspects, nevertheless wound up with the exhortation, "Poverty must go", and gave evidence throughout his book of being ethically motivated. In a later article [95] he affirmed that immediate economic welfare is studied through the distribution of income, and that real or psychic income, corrected for changes in the purchasing power of money, is the goal of the income statistician.

Streightoff's purpose, unless it was the passive goal of showing that data on which a distribution of income should be based were lacking, is difficult to detect. If his interest was primarily statistical, then he must also be credited with unusual conservatism. On the other hand, what he actually achieved (not what he might have done, had pertinent data been plentiful) points to the conclusion that his interest was primarily with the social problems involved. He outlined a threefold utility of income statistics: (1) for framing social legislation, (2) in assessing certain kinds of taxes, (3) in influencing individual and public opinion. These aims would be considered too general by modern standards; but they do indicate that Streightoff was thinking about the uses to which he would put income statistics.

Thus far our investigators have not perceived that a knowledge of the distribution of income would be desirable for purposes other than social welfare (apparently used in the consumption sense) and taxation or government finance. Streightoff explicitly stated that he wanted only enough income statistics to make possible analysis of the social questions he raised, not of problems in economics such as wage theory.[96]

[94] *Wealth and Income*, p. 217.

[95] 'Desirable Additions to Statistical Data on Wealth and Income'. *American Economic Review*, VII (supplement), Part I (March 1917), 157–71.

[96] Streightoff, *op. cit.*, pp. 18, 19. He states that more data than he has specified as 'ideal' would be needed were the analysis to include problems outside the realm of his threefold program.

a) *Later studies primarily statistical*

There is no doubt that Macaulay's intentions were primarily statistical; the opening sentence of his section in Volume II of *Income in the United States* sounded this keynote, while in the more general discussion of Volume I the social implications of the findings were carefully eschewed. The National Bureau set out to answer the question, "How is the aggregate income divided among individuals?" [97] It answered this question by stating, in Lorenz curve fashion, that certain percentages of the people received certain percentages of the income. Further, as a consequence of O. W. Knauth's study of incomes above and below $2,000, and Macaulay's temporal analysis of the tail of the income distributions, the National Bureau made the observation that "the net effect of our participation in the war was to diminish somewhat (at least temporarily) the inequality of the distribution of American incomes" and "if we consider the 5 per cent of those gainfully employed who had each year the largest incomes [over $3,000, approximately], we find that their share in the aggregate of personal incomes declined from about 33 per cent in 1913–16 to about 25 per cent in 1918–19." [98] But these are hardly comments on the ethics of income distribution.

The only motive mentioned by King in the case of the 1921 distribution of income has already been noted in the discussion of the 1921 distribution of wealth (see above Sec. I, 2, b). That is, in the manuscript King considered the distribution of income as helpful to reformers and sales managers; a more extensive teleological discussion is wanting.

King's 1928 distribution, in the construction of which he was commissioned by a New York bank, was apparently built up for a specific purpose. Since the design is not revealed, the applicability of King's results can hardly be appraised. The fact that his sponsor was interested primarily in the larger incomes may have justified King in giving "practically no attention . . . to the distribution . . . among persons having incomes below the income tax limit".[99]

[97] Mitchell, King, Macaulay, and Knauth, *op. cit.*, I, 1.
[98] *Ibid.*, pp. 146–7.
[99] Letter from King dated April 4, 1938.

Leven's conversion of this distribution to a family basis was specifically intended to assess the ability of people of different income levels to pay for medical care on (a) an individual basis, (b) a group basis. Thus he properly deducted, from total income being distributed, $5.3 billion of imputed income (from durable consumer goods), while the family income basis is more defensible for this purpose than for certain other uses to which income distributions are put. In the problem of medical care, the family does seem to be the significant unit. Finally, Leven was probably justified, when debating whether to include families supported by others, in deciding that the source of income was "perhaps immaterial".[100] However, a perusal of the resulting publication fails to reveal what use was made of this comprehensive income distribution. Instead, sample income distributions were relied upon to measure the ability to pay for medical care. Finally, there is King's admission that his original distribution, on which Leven based his construction, was mainly an upper-bracket income study; so one may question its applicability to the problem of medical care.

b) *Brookings and N.R.C. study consumption*

As students of the income distribution problem, the Brookings investigators stand out in several respects. Their construction of the distribution of income in 1929 is more fully explained than any preceding distribution; moreover, in lieu of fuller explanations of the earlier attempts, we may conclude that the Brookings estimate is at least as thorough and rigorous; finally, the Brookings project was not confined to the statistical aspects of the problem, but embraced in an unprecedentedly thorough fashion certain implications of the distribution of income. It thus achieved a balance between statistical and social purpose that is singularly lacking in the earlier distributions.

The keynote was sounded in the Foreword: "The purpose of the investigation as a whole is to determine whether the existing distribution of income in the United States among various groups in society tends to impede the efficient functioning of the economic system."[101] Later, the goal sought by constructing a dis-

[100] Reed, *op. cit.*, p. 101n.

[101] Leven, Moulton, and Warburton, *op. cit.*, p. 1.

tribution of income was more specifically stated: "If, therefore, we are to get a picture of the effective consuming capacity of the American people as a whole, and of the allocation of the national income as between consumption expenditures and savings for the development of capital equipment, we must first see the way in which the income of the nation is distributed among families and other income recipients." [102] It is apparent that Leven and his colleagues were interested in two social aspects of the distribution of income: its effect on consumption and on savings. Furthermore, this purpose is assiduously pursued, and with the aid of data additional to the family distribution of income, conclusions are reached concerning these questions. The validity of these conclusions is less important to us than the fact that here was a distribution of income specifically constructed for and actually utilized in studying certain predetermined social and economic problems. In doing this, however, the Brookings Institution was pioneering in only one respect: that of both constructing and interpreting a distribution of income. Interpretations of distributions of income figures by students other than those who compiled them have been frequent, as witness the scores of books in marketing and consumption economics. Finally, in limiting themselves to these two implications of the distribution of income, Leven and his collaborators failed to consider, except in a very general way, the bearing their results had on taxation, velocity of money, law of demand, wage theory, and related knots in economics.

The purpose of Tucker's article in the *Quarterly Journal of Economics* is not far to seek. The general impression is that the writer is striving to justify, or at least to paint in favorable colors, the existing distribution of income in the United States. As stated in the article, however, the reasons for studying income distributions are two: (1) static analysis of the income distribution at a moment of time is "fundamental to any sound analysis of present-day social problems"; (2) "knowledge of how that distribution has changed is essential for any sound judgment concerning the progress of the nation and the merits or defects of the capitalistic system." [103] Tucker adds that, although the size

[102] *Ibid.*, p. 51.
[103] *Op. cit.*, p. 547.

of the national income is important of itself, so is its distribution since the latter leads to class cleavages and determines the rapidity of capital formation. Nothing along the lines indicated by these two explicit purposes is essayed in the body of the article. Tucker was concerned mainly with working out measures of inequality over time, but the reader is left with the general impression that justification of the status quo was also a desideratum.

The National Resources Committee's contribution to our knowledge of the distribution of income is "part of a larger study of the Nation's consumption demands in relation to its productive capacities".[104] This distribution was therefore devised primarily for use in compiling national estimates of consumer expenditures. The Committee points out some of the purposes the distribution of income could serve: "Those concerned with the living standards of the people need more accurate information on the extent to which shortage of income brings poverty damaging to health and happiness. Lawmaking bodies striving to apportion taxes equitably and without damage to the processes of industry need to know what will swell or deplete the streams. Business men require more abundant and reliable data on the probable demand for their products in order to stimulate and meet that demand. Any attempt on the part of Government or business to grapple with basic economic problems must rely heavily on what can be learned of the distribution of income among the various groups of the Nation's consumers." [105] Application of their distribution to these broader social purposes, however, is not essayed, although segments of such an analysis are promised in subsequent publications. In the volume the inequality in the distribution of income is made manifest by comparisons among different tenths of the population, while the discussion of 'The Three Thirds of the Nation' is probably not wholly dissociated from President Franklin D. Roosevelt's remarks concerning the plight of 'one third of the nation'. Such analysis provides good substance for newspaper editorials and discussion of social questions, but it hardly constitutes a scientific presentation of the significance of the existing distribution of income. Final judgment, of course, must await presentation of

[104] National Resources Committee, *op. cit.*, p. 1.
[105] *Ibid.*, p. 1.

further studies now being carried out by the National Resources Committee.

3 STATISTICAL EVALUATION

a) *Early studies largely guesswork*

It is difficult to assess the statistical adequacy of these attempts to construct a frequency distribution of income, for the methods followed were explained only in rough outline, while the computations were usually not shown at all.[108] The earlier attempts are generally less reprehensible in this respect than the later, but the statistics of the latter should logically be less open to criticism. The anomaly of explaining the method when it is obviously makeshift, and concealing it when it is likely to be more sound, probably arises from the fact that the statistical manipulations involved in these later constructions were so detailed and complicated that their description was impracticable. However, complexity is no defense of unexplained methodology, and indeed may constitute a criticism thereof, for the reason that (a) this complexity may be an attempt to gloss over inherently simple but fundamentally unsound methods, or (b) such complexity, even though genuine, may have deluded the investigator himself. In any case, the student is left with a dissatisfied feeling after perusing an unexplained statistical construction, especially one whose figures are carried out to several decimal places, thus conveying a sense of accuracy unwarranted by the facts.

Spahr's method, although more fully explained and less complex than that of the later investigators, is vulnerable at several points. He based the distribution of family incomes on the distribution of estates, so that he had to bridge not only the previously mentioned gap between estates and families, but also the more hazardous interstice between wealth and income. Equal wealth does not make for equal income, nor is all income derived from physical wealth; much income springs from human skills, knowledge, experience and, in general, labor. Furthermore, the proportion (two-fifths) of total income which he assigns to capital was not only a guess but also probably an exaggeration. Even today, with our economy more heavily mechanized than in

[108] The National Resources Committee's distribution is a possible exception.

1890, salaries and wages constitute two-thirds (not three-fifths) of the net national product, while the return to capital, even including entrepreneurial withdrawals, amounts to no more than one-third (not two-fifths).[107] In addition, Spahr's use of the alleged function—propagated by Paul Leroy-Beaulieu [108]—relating rent paid to income received, is of doubtful justification. According to A. L. Bowley in his review of Streightoff's *The Distribution of Incomes in the United States,* experience in England "shows that the relation between rent and income is variable and complex." [109] Finally, other lacunae in Spahr's analysis have already been indicated: e.g., his percentage return on the wealth holdings of each income class and his derivation of the average labor income of the well-to-do and wealthy classes.

Streightoff did not get far enough in his construction of a distribution of income to warrant criticizing his method. However, his proposals for ascertaining (a) farmers' expenses and (b) paid and imputed rent of urban dwellers are open to debate. Bowley criticized both these suggestions on the ground that (a) farmers do not know their expenses and (b) the rent-income function is variable and complex. The second criticism stresses a point too often neglected by statisticians, but the first is an unnecessarily pessimistic commentary on the knowledge and aptitude of farmers. Although it is true that they would have the same trouble, perhaps somewhat augmented, that a business man has in preparing his income tax blanks, yet the problem seems not insuperable, and once the farmers were trained to keep elementary accounts, the information Streightoff proposes gathering would be invaluable in distributing farmers according to the size of their net income.

Streightoff also lists his criteria of ideal income statistics: (a) urban incomes segregated from rural, (b) incomes adjusted for standard of living and purchasing power of money, (c) small class intervals, (d) incomes classified according to source (from property, labor, etc.), (e) occupation, residence, and race of income recipients, (f) complete returns from every gainfully

[107] Department of Commerce, *Income in the United States, 1929–37* (Washington, 1938), p. 22.

[108] *Répartition des Richesses* (Paris, 1897).

[109] *Economic Journal*, XXIII (1913), 425–7.

employed. This is quite an order, but it is significant that he failed to give a definition of what shall be considered income, he did not specify whether he would use the individual or the family as his income recipient, and in calling for complete returns from every gainfully employed he ignored the many income recipients without gainful occupation.[110]

b) *King and Macaulay weak on explanations*

The methods employed by King in his pioneer work on *Wealth and Income* were unfortunately insufficiently explained to make it possible to assess their propriety. He was probably justified in considering the Wisconsin distribution as representative, at least for the middle section of his composite distribution, while his use of earnings data and income tax statistics for the lower and upper ranges respectively would seem logical. In arguing for a distribution based on families rather than individuals, he failed to see that for some purposes the former is preferable, and for others, the latter; but this is less a criticism of what he has done than of what he has left undone. The reviews of King's first attempt seem unnecessarily critical. G. P. Watkins dismissed it with the charge that King's "faculty of statistical analysis" was inadequate,[111] while A. A. Young in his more dispassionate review concluded that the method King used in estimating the aggregate annual product (which was distributed among the 28,500,000 families) "must have involved a large amount of conjecture".[112] He based this conclusion on the allegation that the federal income tax returns for 1913 showed that (a) King's scheme of distribution was "very much awry", or (b) his estimate of aggregate income was "very much too large", or (c) the federal government got only three-fifths of the income it was entitled to under the law. Young did not make clear how he arrived at this criticism. While the whole discussion lies outside the proper realm of this paper, it may be observed that the subsequent National Bureau

[110] Since in subsequent discussion he recognized ownership of property and rights of private property (e.g., gifts and inheritance) as sources of income. Streightoff probably would not have been long in discovering, in an actual calculation, this last-mentioned oversight.

[111] *American Economic Review*, VI (1916), 443.

[112] *Quarterly Journal of Economics*, XXX (1916), 358n.

estimate of the national income in 1910 was even larger than King's figure.[113]

The method employed by Macaulay in 1921 was also too sketchy to allow much criticism. His concept of the personal income recipient was not clear. Apparently it fails to comprehend non-gainfully employed persons with income, yet he significantly fails to say so. In 1929 Leven estimated there were 2,000,000 income recipients without a gainful occupation,[114] so this may be a significant confusion in Macaulay's analysis. Furthermore, Macaulay's wholesale adjustment of income tax data for underreporting and negative incomes, and his smoothing of the final distribution curve raise several doubts in the reader's mind, especially since these adjustments are largely unexplained. Perhaps for these reasons, Bowley suggested that Macaulay might well have postponed publication of his estimates, or at least have buried them under his mathematics. "It is inadvisable that very doubtful estimates should be given currency. . . . Statisticians are sometimes inclined to let their desire to obtain a complete statement overcome their knowledge of the insufficiency of materials," the English statistician commented.[115]

King's 1921 distribution, as the description suggests, was devised by means of one arbitrary assumption after another. The results simply cannot be given the credibility demanded by the detail in which they are presented. The passage from gross to net farm income by means of the crop reporter sample is unconvincing, even though King does strive to correct for the lack of randomness in the sample by means of an arbitrary adjustment. The use of income tax data to adjust the distribution for incomes above $2,000 is broader than any other writer has dared make. Because of exemptions and credits to income, it is generally recognized that these income tax data are unsuited for this adjustment below some such level as $5,000. In fact, the National

[113] King's estimate was $30.5 billion, while the National Bureau put the national income produced in 1910 at $31.8 billion (Mitchell, King, Macaulay, and Knauth, *op. cit.*, I, 13).

[114] Leven, *op. cit.*, p. 186. In fact, in adjusting Macaulay's distribution to 1929 conditions, Leven felt constrained to supplement it with those "income recipients who were not gainfully employed" (*ibid.*, pp. 177, 178).

[115] Review of National Bureau's *Income in the United States*, *Quarterly Journal of Economics*, XXXVII (1923), 510–17.

Resources Committee used these data for adjusting incomes only above the $7,500 level. Furthermore, even if it is agreed that negative incomes are a legitimate constituent of the desired distribution, King's method of estimating these negative incomes by extending the curve of positive incomes, freehand, into the negative side of the graph can hardly be condoned. Finally, one wonders why King further confused the issue by the employment of at least four different concepts of income. Certainly the more refined concepts made the resulting synthesis not only more arbitrary and fictitious, but also more vulnerable to criticism against the general policy of considering unrealized capital gains and imputed interest as income. The chief argument in its defense, in the present instance, is that King evaded the question of which concept of income to use, by constructing distributions based on all the different concepts and letting the reader take his choice. Such a procedure may be statistically commendable, but it does further confuse an already complicated mosaic.

King's distribution for 1928 obviously cannot be evaluated, and Leven's conversion of it into a family distribution merits only passing review. Leven himself admitted "that the estimates are extremely rough and only tentative" [116] while the fact that he elaborated his procedure in the subsequent Brookings inquiry makes an appraisal of this earlier work redundant. The many assumptions involved and the absence of clarity in his method have been indicated.

c) *Leven draws several criticisms*

Passing over King's temporal analysis of income distributions based solely on *Statistics of Income* tables and the Census estimate for gainfully employed, we reach the 1929 distribution of income by the Brookings Institution. This study, in sharp contrast to its predecessors, is replete with details concerning the methods employed and calculations made: yet certain significant explanations are missing. Information on how the 'under $0' class was estimated is meager,[117] and, as Burns pointed out, no

[116] Reed, *op. cit.*, p. 101.

[117] It seems designed to approximate the figure in *Statistics of Income for 1929* (Washington, 1931), p. 11, for "Loss from sale of real estate, stocks, bonds, etc., other than reported for tax credit". This is a mere surmise, however.

explanation is made as to how the independent estimate of personal incomes in 1929 was broken up into three distributions: (a) incomes of heads of families of two or more persons, (b) supplementary income recipients in families having more than one income, (c) incomes of unattached individuals living alone.[118] The inclusion of realized capital gains and losses in income has been criticized on the grounds that since the former swell the number of families in the upper income brackets and the latter presumably dominate the negative income class the distribution's utility in a study of savings is impaired.[119] Moreover, the synthetic families (artificial compounds of breadwinners and supplementary earners) used as the unit of the income recipient seem less defensible than existing or economic families, or ammains. Certainly they introduce an unreality into the distribution that makes one wonder just what the final figures represent. Burns further characterized the personal income distribution underlying the family distribution as a patchwork based on scanty data and some dubious statistical devices. Among the latter he stressed the conversion of actual gross farm incomes into net farm income equivalents by means of a scanty sample and rank, instead of identical farms, correlation.[120] Furthermore,

[118] Nor does Leven explain how the distribution of families with two or more income recipients was divided into two distributions of (1) families having only two income recipients, and (2) families having more than two income recipients: etc.

[119] Burns, *op. cit.*, p. 495. This effect on the distribution was admitted by Leven, Moulton, and Warburton, *op. cit.*, p. 57, and defended on the ground that such gains and losses "must be included in the income received by individuals if we are to discuss intelligently the flow of income from individuals into consumption and investment channels" (*ibid.*, p. 13). The error arises, as Leven confessed (p. 131), from the impossibility of segregating capital gains considered as income from capital gains considered as capital. This psychic difficulty would suggest omitting realized as well as unrealized capital gains and losses from income, when studying consumption and savings.

[120] This raises an interesting statistical problem. The correlation of gross with net farm incomes, in the samples, was made by first arranging the gross and net farm incomes in separate arrays, from low to high, and then by associating a given gross income with that net income occupying the corresponding rank in the array. That is, low gross incomes were associated with low net incomes, and so on. This method of correlation gives a higher coefficient than that in which a given gross income is associated with the corresponding net income for the same farm. It also, as Leven observes (p. 200n), has the effect of increasing the slope of the line of regression of the net farm income on the gross farm income, thereby swinging this

the correction of incomes above the $5,000 level for underreporting and evasion is not clear. Leven states that the estimated number of income tax returns of persons reporting income from business and professions has been raised 65 per cent. This percentage is based on Macaulay's experience with the 1918 data, and on Leven's own survey, *The Incomes of Physicians*. He does not say whether the same percentage was used in correcting each income class. Nor does he make clear whether it is the total income of these reporting persons that is increased, or just their income from business and profession. Finally, several of the numerous assumptions inherent in Leven's calculations have already been cited; some may be empirically vulnerable, while all are certainly open to debate. But regardless of their general validity, the question arises whether the use of such algebraic relations, e.g., between occupational and total income or between gross and net farm income, does not conceal basic differences between the incomes of individuals or families—differences it is the purpose of a distribution of income to reveal. It may be true that, in general, net farm income is a certain function of gross farm income; but the fact that this function varies from farmer to farmer within a given gross-income range is one of the many reasons for inequality in farm incomes. The assumption that this function is constant for a given income class would have the effect of concealing important inequalities.

Statistical evaluation of Tucker's article is not appropriate

line counter-clockwise about the mean value. By increasing the slope of the regression line Leven in effect obtained lower values for net farm incomes in the lower income brackets than would have been obtained by identical-farms correlation. Leven argues in defense of this procedure that he did not wish to find the net income for the same farmer for whom gross income was known, but rather that he wished to reconstruct the distribution of net incomes for the entire group of farmers. There is clearly some point to his argument, since the use of a regression equation based on identical-farms correlation to estimate the distribution of net income inevitably tends to yield a distribution less dispersed than the 'true' one (cf. Part Two, section 1 of discussion by Milton Friedman). However, it is doubtful whether this argument fully justifies Leven's procedure, in view of the difficulties of attaching any clear and unambiguous meaning to it. Moreover, the higher correlation coefficient obtained by arranging the items in arrays does not increase the representativeness of the crop reporter sample, nor does it correct for the fact that the samples used to derive the relationship between gross and net were admittedly limited in large part to the more successful and better-remunerated farmers.

since he failed to construct a complete distribution of income. This does not deny the fact, however, that certain of his analytical devices seem to be misleading, while his contention that changes in the distribution of taxpayers' income indicate changes in the complete distribution is subject to considerable doubt. A moment's reflection will show that regardless of what happens to the upper income distribution, a shift in the location of the modal income, or a flattening out of the lower portion of the income distribution—to mention only two possibilities—would significantly alter the effective inequality of incomes.

d) *National Resources Committee*

For statistical adequacy the distribution offered by the National Resources Committee leads the list of American distributions of incomes. This is not to say that Miss Kneeland and her staff have constructed a 'correct' distribution of income in any absolute sense, but rather that they have come nearer the desired goal than any of their predecessors. In all fairness, it should be added that credit for this achievement is not necessarily due to any technical or statistical superiority of Miss Kneeland and her colleagues. Although they are undoubtedly competent statisticians, it must be admitted that they had at their disposal better and more abundant original data on which to base their distribution of income than any of their American predecessors in this field. Credit is due them mainly for exploiting rather fully what source material was available. In addition, they deserve commendation for explaining not only in some detail but also with laudable clarity the methods and assumptions used in passing from the sample data to the global distribution. In this respect, too, they stand out from among their predecessors.[121]

The major weaknesses of the National Resources Committee's distribution admittedly center on the use of the income tax data and the handling of the relief item; in addition, such points as the exclusion of institutional residents from the final distribu-

[121] This does not deny the fact, however, that the description in *Consumer Incomes in the United States* of the adjustments made by means of data from income tax returns still leaves the reader under somewhat of a cloud. A fuller explanation of these adjustments is necessary. Such an explanation has been prepared by Enid Baird and Selma Fine and appears below in Part Three.

tion come in for incidental criticism. These first two considerations become particularly important when it is remembered that inequality in the distribution is profoundly affected by seemingly minor adjustments made at either extremity of the frequency table. The third point is relatively unimportant. Since institutional residents do not constitute a spending unit in the same sense that families and single individuals do, the Committee is quite justified in omitting them from the distribution in the study of consumption. But the Committee, in the volume under consideration, uses the final distribution to point out differences in the welfare of various segments of the population; for this purpose, the argument for excluding institutional residents is not so clear. Inclusion of the institutional residents would not only alter slightly the comparison of the three-thirds of the nation but also make the comparison more realistic.[122]

On the problem of relief outlays, little comment is necessary. This particular question is peculiar to the present distribution: no previous investigator was faced with the necessity of distributing nearly a billion dollars in direct relief among the family and individual income recipients. In addition, the Committee itself confesses: "The methods used in adding the value of direct relief to these income distributions were necessarily crude, and involved various arbitrary assumptions based upon very fragmentary evidence from available relief studies." [123] How radically the distribution would have been affected by different assumptions and methods is problematical.

The adjustment of the distribution by data from income tax returns, on the other hand, warrants more critical inspection. Before utilizing the federal income tax data it was necessary to combine the incomes of husbands and wives making separate returns. This was effected by means of the general assumption that "at the high income levels husbands and wives making separate returns would endeavor to divide the family income as evenly as possible in order to avoid the surtax charges".[124] Some

[122] In Chart 6 on p. 8 of the report, the lower third single men and women would be increased approximately 3½ figures (or symbols), and the middle third men about ½ a figure, were the institutional residents as classified in Table 16 on p. 32 taken into consideration.

[123] National Resources Committee, *op. cit.*, p. 65.

[124] *Ibid.*, p. 83.

such assumption is admittedly necessary, but the present one seems to do violence to our sense of expectations. It is hard to believe, for example, that even the majority of the so-called 'economic royalists' share their properties and incomes evenly with their wives. Instead of pairing women with high incomes against men with high incomes, it would seem just as fitting to pair high-income women with medium-income men. The problem, however, is admittedly difficult; any system of pairing would have to be arbitrary.

The correction for nonreporting and understatement of incomes admittedly is likewise exceptionally artificial. Just why it was decided, for instance, to increase the number of families in the $5,000 to $10,000 income class 25 per cent is hard to perceive. And since no referable explanation is given, one is forced to conclude that it was largely 'drawn out of the air'. Also, how was the decision reached to increase the aggregate income of the same class 15 per cent? It would seem that if such corrections are going to be made, some sort of basis for selecting the given percentages, other than a vague reference to "tentative estimates advanced by several authorities", should be indicated. Otherwise the careful reader is left unconvinced, while the untrained reader is given a sense of accuracy in the adjustment that is belied by the facts.[125]

Finally, it is unfortunate that the passage from statutory net income to economic income as defined in the study could not be effected more satisfactorily. Because only preliminary tables of certain 1935 income tax data were available, the National Resources Committee was forced to carry through this transition by means of at least two arbitrary assumptions. The first was that the necessary additions to statutory net income, (i.e., for net capital losses, contributions, taxes paid, interest paid, and tax-exempt interest received) and deductions from statutory net income (i.e., for capital gains) were distributed among the various groups of return (joint, separate, etc.) at each income level above $5,000 "according to the proportions of aggregate net income [statutory]

[125] *Ibid.*, p. 84n. In this connection the reasoning underlying the following footnote is interesting: "The sequence of the adjustments for nonreporting and understatement [the former was made first, and the latter second] implies that families added to the distribution to allow for nonreporting would have understated their incomes to the same extent as did the families that actually filed income tax returns."

received by each group at the various levels".[126] Second, within each income class the combined additions and deductions required to pass from statutory net income to economic income were apparently divided evenly among the income recipients in each class. Both assumptions are far from obvious, and their use necessarily attenuates the reliability of the resulting distribution.

The distributions constructed by the Brookings Institution for 1929 and by the National Resources Committee for 1935–36 are the most satisfactory thus far presented, all things considered. In the accompanying outline, certain respects in which they differ are pointed out, in order to illustrate some of the many decisions the investigator must make.

BROOKINGS INSTITUTION	NATIONAL RESOURCES COMMITTEE
Nature of Sample	
The lower ranges are based on a composite of many small samples for varying years and groups, adjusted to 1929 conditions.	The lower ranges are based on the Consumer Purchases Study of 300,000 families representing various regions and groups in 1935–36.
The income classes above the $5,000 level are based on *Statistics of Income* data.	Above the $7,500 level the distribution is based on *Statistics of Income* data.
Differences in Elemental Definitions	
The definition of income provides for the inclusion of capital gains and losses. Hence, the negative income class became a significant part of the distribution. Moreover, supplementary incomes (as opposed to earnings) are included.[127]	The definition of income excludes capital gains and losses, except in lower income classes on goods exchanged within the year. No negative income class was segregated.

[126] *Ibid.*, p. 82; Part Three, Sec. IV, especially note 8.

[127] Apparently the same items of non-money income—farm produce consumed on the farm and imputed rent on owned houses—are included in both distributions. There may be slight differences in details, however.

Differences in Elemental Definitions—Cont.

A slight understatement occurred because of the inclusion for income levels above $7,500 of supplementary earnings rather than supplementary incomes.[127] At lower income levels, supplementary *incomes* were included.

The definition of family refers to census-biologic families, in which blood relationship is the primary attribute. The actual joining of the supplementary income recipients to the main breadwinner is partly a matter of chance, and the 'families' are more or less 'compounds'.[128]

The definition of family refers to economic families, *i.e.*, living under one roof and having a common or pooled income. Some arbitrariness arises in the pairing of husband and wife in the upper income classes, and in the exclusion of self-supporters from the family.[128]

Differences in Procedure

Occupational income was derived, made to jibe with the Department of Commerce estimate, then supplemented with (1) other income of the gainfully employed, (2) that of those without gainful occupation.

The incomes of farmers were estimated separately by means of Department of Agriculture figures and the crop reporter sample.

Up to the $15,000 level, families were synthetically built up by joining supplementary income recipients to main breadwinners.

The nature of the sample made it possible to pass directly to total income (except a small item of supplementary unearned incomes), which figure happened to jibe tolerably closely with the Department of Commerce total.

Since farmers' incomes were included in the Consumer Purchases sample, the necessity for separate estimation did not arise.

The lower ranges were already on a natural—or existing—family basis, because of sample. Arbitrary pairing was employed to put upper ranges on a family basis.

[128] Neither distribution, of course, avoided certain artificialities inherent in the Census definition of family.

Correction of Income Tax Data

Underreporting and evasion in the incomes of $5,000 and over were set at 65 per cent; the estimated number of income tax returns for business and professional incomes was increased by that percentage.

Corrections for nonreporting follow: increase the number of families and aggregate income in the $5,000–$10,000 class 25 per cent; in the $10,000–$15,000 class 15 per cent; in the $15,000–$20,000 class 5 per cent.

Corrections for understatement: increase the aggregate income of families in the $5,000–$10,000 class 15 per cent; in the $10,000–$20,000 class 15 per cent; in the $20,000–$25,000 class 10 per cent; in the $25,000–$50,000 class 5 per cent. The overall estimate of understatement equaled 10 per cent, and was made after the nonreporting correction had been introduced.

This correction was applied to the distribution of occupational income.

These corrections were applied to the distribution of total money income, before the addition of the supplementary earners.

Inclusiveness of Final Distribution

Constituents of institutional groups are included in the final distribution as unattached individuals in the category of spending units.

Institutional residents are not included in the final distribution, although their incomes are estimated and presented separately.

4 RECAPITULATION OF INCOME DISTRIBUTION STUDIES

The accompanying outline presents in summary form the more important characteristics of the studies of the distribution of income discussed in this section.

NAME	OPUS	DATE OF DISTRIBUTION	BASIS OF ESTIMATE	UNIT	PURPOSE, REMARKS
C. B. Spahr	1896	1890	Wealth holdings and common observations	Census family	Social - taxation problem analyzed.
F. H. Streightoff	1912	1904	Labor earnings	Occupied persons	To show necessary data lacking; included income from male labor alone.
W. I. King	1915	1910	Labor earnings Wisconsin income tax data, federal data on the rich	Census family	Primarily statistical.
F. R. Macaulay	1921	1918	Wage samples and federal income tax data	Income recipients	Statistical; included negative and imputed income.
W. I. King	Unpub.	1921	Federal income tax data, wage and budget samples	Income recipients	Statistical; for New York bank.
	Unpub.	1928			
Maurice Leven	1933	1928	Converted King's 1928 estimate	Family (synthetic)	To assess people's ability to pay for medical care: tentative conversion to families.
Maurice Leven	1934	1929	Wage and farm samples, federal income tax data	Family (synthetic)	Consuming capacity and saving power; also statistically ambitious.
National Resources Committee	1938	1935, 1936	Study of Consumer Purchases and income tax data	Economic family	National consumption estimates: best distribution yet constructed.

III Conclusion

One conclusion from the foregoing review seems inescapable: statisticians and economists have been striving valiantly to fill the persistent demand in this country over the last generation or two for statistical evidence concerning the distribution of wealth

and income. Evidence of their efforts began to appear some two score years ago. Since then there has been a steady flow of published attempts to distribute by size the wealth holdings and incomes of the people in this country.

The data and methods used and the general aims held by these investigators were indicated in the preceding pages, and an attempt was made to evaluate the statistical adequacy of these distributions. Unfortunately, their evalution must be in relative rather than absolute terms. A relative appraisal of the adequacy of these distributions leads to extensive differences of opinion, a point of which the readers of this paper probably need not be reminded. But even sharper differences of opinion would arise were one to assess in absolute terms the degree to which a given distribution of wealth or income was adequate for the purposes for which it was intended and used. Probably some would aver, in good faith, that all these published distributions were totally inadequate. On the other hand, a few might insist, with more temerity than propriety, that a given distribution was entirely adequate. Although the persons at the former extreme would probably outnumber those at the other, the majority of qualified observers would likely fall into a middle class of those who hold that the existing distributions give a rough idea of the actual distribution, but that they are too crude and inaccurate to allow measuring temporal and spatial differences in the inequality of distribution—differences that must be known if changes in relative welfare of different social groups, in tax burden and taxable capacity, in the volume of individual savings and in the pattern of consumer demand are to be analyzed. Furthermore, this temperate group of observers would perhaps question whether existing distributions give a true picture of the relative welfare of the different strata in society, even at a given moment; or make possible a thorough analysis of the existing patterns of consumer demand and individual savings.

A crucial question emerges from these considerations. Why, in view of this pressing and widely felt need for accurate statistical information on the distribution of wealth and income, do we not have better and more adequate data? Why has there been this time lag between the realization of a need and its satisfaction?

That American statisticians have been unusually inept, or that they have been unwilling to give reign to the imagination can hardly be cited as reasons why their efforts in this field have met with such limited success. In fact, the reverse is probably nearer the truth. Many investigators have been too ambitious and have overstepped the limitations of their data in striving to construct statistical pictures of the distribution of wealth and income.

The reasons why the statisticians have been thus thwarted fall into three general classes. On the one hand we have the psychological factors which lead a person to consider his own income and wealth a secret even though he may be among the front ranks of those clamoring for statistical information on wealth and income distribution. Moreover, this spirit is probably fostered by democratic institutions that aim to exalt free private enterprise, individualism and personal liberty—all with a minimum of government interference. It is also undoubtedly encouraged by the fear that any personal income and wealth information will be used for taxation purposes. Whatever the psychological, social, and political factors contributing to this attitude of reservation, its existence is strongly attested by those who have had the occasion to attempt, through field surveys or otherwise, to procure wealth and income statistics from a considerable number of persons. Hesitancy about answering questions on income and wealth is more pronounced in the upper economic classes; but some maintain that persons in the lower strata, although more willing to divulge the desired information, nevertheless commonly misstate their incomes, either through ignorance or design. In fact, it has become almost an axiom that the adequacy of the response to a questionnaire or field survey varies inversely with the number of questions on such personal matters as income and wealth.

The second group of reasons accounting for the failure of statisticians to meet this demand for statistical information on the distribution of wealth and income lies nearer their own doorsteps. It is the failure to reach satisfactory definitions of wealth, of income, and of family or whatever wealth-holding and income-receiving unit is being employed. Irving Fisher has observed that there is hardly a *communis opinio* among economists on the definition of income.[129] Simon Kuznets has indicated some of the

[129] 'Income' *Encyclopedia of the Social Sciences*, VII, 622–5.

obstacles in the way of selecting a usable and meaningful definition of wealth.[130] The idiosyncrasies of the Census definition of family are well known. That the National Resources Committee, to cite only one example, felt constrained to employ a somewhat different definition of family is itself evidence that the Census concept is not wholly appropriate for income distributions. Macaulay tried to avoid some of these difficulties by using individuals rather than families. Whether such procedure eludes more difficulties than it raises is open to question. If the individual basis is used, then King's further refinement to 'ammain' warrants consideration in certain cases.

A third possible reason why attempts to construct distributions of wealth and income have been relatively unsuccessful is that, in addition to the meagreness of primary data, functionally related series are also scarce. There seem to be few economic series sufficiently closely and simply related to income and wealth to make it possible to derive the latter from the former. The attempts repeatedly made in this direction have thus far not been attended by particular success. Efforts to derive the distribution of income from the distribution of rents are one example; the method of estimating wealth holdings by capitalizing income is another.

From this brief discussion it would seem that several obstacles must be overcome if a satisfactory distribution of income, for example, is to be constructed. The first is a decision as to the purpose of the distribution. Once that has been made, a suitable definition of income and the selection and definition of the income-receiving unit (family or individual or ammain) must be made. After this underbrush has been cleared away two courses are open to the statistician; these may be considered either as alternatives or as complements. The first is the direct procurement of statistical information, either for a sample or the universe, on the incomes of the (say) families. This, in turn, may require either a remolding of the people's mores in the direction of inducing them to divulge more freely and accurately the desired information concerning their incomes, or legislative enactment making obligatory reporting of incomes to the Census, or possibly both.[131] The second course, which may be employed

[130] *Volume Two*, Part One.

[131] Even the statutory requirement would have its limitations. Chances of so

either as a substitute for or a complement to the first, is to ascertain the distribution of rents for relatively homogeneous groups, derive the rent-income function for each such group, and pass therefrom to the distribution of incomes. Weighting and summation would then give the global distribution of incomes.

It would be hazardous to conjecture precisely when statisticians will succeed in overcoming the present obstacles to really adequate distributions of income and wealth for this country. Because of renewed positive interest of public as well as private research organizations in the problem, one may expect in the near future not only better distributions but also distributions adequate for statistical induction.

Bibliography

This compilation is limited in general to American writings on the subject of this report and is not intended to be co-extensive with the references cited in footnotes. The principal omissions are the source material on income distribution now being issued (usually in mimeographed form) by various government agencies, other sample income and earnings surveys which are not only extremely numerous but are also listed in other compilations, and reports and bulletins preliminary to and superseded by later and more comprehensive publications. Those references concerned with the problem of measuring wealth or income inequalities are included primarily because they consider a neglected and crucial issue. Titles of certain journals have been abbreviated, as follows:

P.A.S.A. Publications of the American Statistical Association (title for Volumes I–XVII)
J.A.S.A. Journal of the American Statistical Association (title for Volumes XVIII to date)
A.E.R. American Economic Review
Q.J.E. Quarterly Journal of Economics
J.P.E. Journal of Political Economy

changing the people's mores that understatement and overstatement would be avoided, even under oath, are slight. Therefore, intensive study, with the aid of skilled enumerators, of the correction factors needed to eliminate such bias is indicated.

Books

Max Ascoli and Fritz Lehmann, *Political and Economic Democracy* (Norton, 1937), especially Ch. IX, 'Distribution of Wealth' by Lehmann

W. A. Berridge, E. A. Winslow, and R. A. Flynn, *Purchasing Power of the Consumer* (J. W. Thompson Co. prize essay, 1925); especially good for bibliography on factory workers' incomes

H. F. Clark *et al., Life Earnings in Selected Occupations in the United States* (Harper, 1937)

M. A. Copeland, 'The National Income and Its Distribution' in *Recent Economic Changes* (National Bureau of Economic Research, 1929) II, 757–839

W. L. Crum, 'The Distribution of Wealth', *Harvard Business Research Studies,* No. 13 (1935); a distribution among the well-to-do, based on federal estate tax returns

R. R. Doane, *The Measurement of American Wealth* (Harper, 1933)

A. J. Ferris, *Pauperizing the Rich* (Philadelphia: T. S. Leach, 1899)

Harvard Law School, *Earnings of Lawyers* (West Publishing Co., Docket No. 3:3004, March 1936)

Yandell Henderson and M. R. Davie, *Income and Living Costs of a University Faculty* (Yale University Press, 1928)

Income Record (New York: American News Co., 1865)

Income Tax of the Residents of Philadelphia (Philadelphia: John Trenwith, 1867)

Income Tax of the Residents of Philadelphia and Bucks County (Philadelphia, 1865)

W. R. Ingalls, *Current Economic Affairs* (York, Pa.: Merlin, 1924)

Wealth and Income of the American People (2d ed.; York, Pa.: Merlin, 1924); two chapters on division of wealth and income, by size

E. A. Keller, *A Study of the Physical Assets, Sometimes Called Wealth, of the United States, 1922–1933* (University of Notre Dame, 1939); an extrapolation of Federal Trade Commission figures for 1922, this study also denounces certain popular beliefs concerning inequality in the distribution of wealth

W. I. King, *The National Income and Its Purchasing Power* (National Bureau of Economic Research, 1930)

Wealth and Income of the People of the United States (Macmillan, 1915)

Simon Kuznets, 'The Measurement of National Wealth', *Studies in Income and Wealth* (National Bureau of Economic Research, 1938), II, 1–61

Maurice Leven, H. G. Moulton, and Clark Warburton, *America's Capacity to Consume* (Brookings Institution, 1934)

Maurice Leven, *The Incomes of Physicians* (Committee on the Costs of Medical Care, University of Chicago Press, 1932)

'Note on the Distribution of Income' in L. S. Reed, *The Ability to Pay for Medical Care*, Appendix A, pp. 99–101

M. O. Lorenz, *The Distribution of Wealth in Six Wisconsin Counties* (unpublished ms.), cited in King, *Wealth and Income*

Richmond Mayo-Smith, *Science of Statistics* (Macmillan, 1899). 2 vols., see especially Vol. II *(Statistics and Economics)*, Ch. XIII, entitled 'Social Distribution'

W. C. Mitchell, W. I. King, F. R. Macaulay, and O. W. Knauth, *Income in the United States* (National Bureau of Economic Research, Vol. I, 1921, summary; Vol. II, 1922, details)

Scott Nearing, *Income* (Macmillan, 1915): a study of individual earnings

L. S. Reed, *The Ability to Pay for Medical Care* (Committee on the Costs of Medical Care, University of Chicago Press, 1933)

C. B. Spahr, *The Present Distribution of Wealth in the United States* (Crowell, 1896)

F. H. Streightoff, *The Distribution of Incomes in the United States* (Columbia University Studies, Vol. LII, No. 2, 1912)

Time, Inc., *Markets by Income* (New York, 1932); a marketing study of Appleton, Wis., based on state income tax returns

H. M. Trumbower, *The Distribution of Income in Wisconsin* (unpublished ms.), cited in King, *Wealth and Income*

Twentieth Century Fund, *Studies in Current Tax Problems* (New York, 1937), see especially article by Mabel Newcomer, 'Estimate of the Tax Burden on Different Income Classes', pp. 1–52; the income classes employed are hypothetical

Reuben Vose, *Wealth of the World Displayed* (New York: privately printed by Mr. Vose, 1859); gives names of, and some gossip concerning, the wealthy men in this country and abroad; alphabetical index of names

D. B. Yntema, *Measures of the Inequality in the Personal Distribution of Wealth and Income* (University of Michigan Ph.D. thesis, unpublished, Ann Arbor, 1931); an excellent mathematical discussion of a difficult and important matter: What is the meaning of inequality? How can it be measured?

Periodicals

Louis Bader, 'The American Family Income and Prosperity', *J.A.S.A.*, XXVIII (1933), 303–11

J. D. Black, 'Farm Business Surveys as Sources of Data on Agricultural Income', *Review of Economic Statistics*, X (1928), 174–81

A. L. Bowley, Review of Mitchell *et al.*, *Income in the United States*, *Q.J.E.*, XXXVIII (1923), 510–17

Review of Streightoff, *Distribution of Incomes*, *Economic Journal*, XXIII (1913), 425–7

A. F. Burns, 'The Brookings Inquiry', *Q.J.E.*, L (1937), 476–523

T. N. Carver, 'The Meaning of Economic Equality', *Q.J.E.*, XXXIX (1925), 473–5; an outline of the various attributes in terms of which economic equality may be defined

W. L. Crum, 'Individual Shares in the National Income', *Review of Economic Statistics*, XVII (1935), 116–30

'On the Alleged Concentration of Economic Power'; G. C. Means, 'A Reply'; and Crum, 'A Rejoinder', *A.E.R.*, XXIV (1934), 69–88

'Regional Diversity of Income Distribution', *American Journal of Sociology*, XLII (1936), 215–25; a discussion of regional differences in inequality, based primarily on income tax statistics

R. R. Doane, 'Summary of the Evidence on the National Wealth and Its Increasing Diffusion', *Annalist*, July 26, 1935, pp. 115–8

'An Accurate National Wealth Census: Statistical and Other Limitations', *Annalist*, Aug. 2, 1935, p. 158

'Tax Payments as an Aid to More Exact Measurement of Wealth Distribution', *Annalist*, Aug. 9, 1935, pp. 189, 214

'Changes in the Distribution of Wealth Since 1880; Greater Diffusion Shown', *Annalist*, Aug. 16, 1935, pp. 222–4

'Rejoinder', *Annalist*, Aug. 30, 1935, pp. 292, 293, 312

'The Geographic Distribution of the Physical Wealth in the United States', *Annalist*, Nov. 15, 1935, pp. 676–9

'Property Ownership by States: Security Holdings, Insurance Equities, etc.', *Annalist*, Dec. 20, 1935, pp. 844–6

'The Division of the National Wealth Between Farm and Non-Farm Property', *Annalist*, Jan. 31, 1936, pp. 196, 197

'Statistical Bases for National Wealth Estimates', *Annalist*, March 27, 1936, p. 478

'Distribution of Corporate, Individual and Public Debts and Equities, 1932', *Annalist*, May 15, 1936, pp. 718, 719, 725

J. F. Ebersole, S. S. Burr, and G. M. Peterson, 'Income Forecasting by the Use of *Statistics of Income* Data', *Review of Economic Statistics*, XI (1929), 171–96, corporations; XII (1930), 39–47, corporations; and XII (1930), 59–76, individuals

R. C. Epstein, Review of Crum's *Distribution of Wealth*, *J.A.S.A.*, XXXI (1936), 770; "The sample is severely truncated."

L. F. Garey, 'Wealth Accumulation by Farmers', *Social Forces*, XI (1932), 120–7; based on a survey of a Minnesota county

J. A. Hill, 'The Civil War Income Tax', *Q.J.E.*, VIII (1894), 416–52, 491–8

G. K. Holmes, 'The Concentration of Wealth', *Political Science Quarterly*, VIII (1893), 589–600

'Measures of Distribution', *P.A.S.A.*, III (1892–3), 141–57

'Measurement of Concentration of Wealth', *P.A.S.A.*, IX (1904–5), 318, 319

R. H. Jackson, 'Full Text of Memorandum on the National Wealth and Its Distribution', *Annalist*, Aug. 30, 1935, p. 292

N. O. Johnson, 'The Brookings Report on Inequality in Income Distribution', *Q.J.E.*, XLIX (1935), 718–24

'The Pareto Law', *Review of Economic Statistics*, XVIII (1927), 20–6

K. G. Karsten, 'An Index of Incomes', *P.A.S.A.*, XVII (1920), 253–76; a proposed method for rigidly classifying economies according to inequality in their distribution of income and wealth

F. S. Kinder, 'The Effects of Recent Changes in Monetary Standards upon the Distribution of Wealth', *Economic Studies* (American Economic Association), IV (1899), No. 6, 409–99; a treatment of the effect of monetary standards on prices, and of prices on the distribution of income

'Monetary Standards and the Distribution of Wealth', *Political Science Quarterly*, XV (1900), 351–5

W. I. King, 'Desirable Additions to Statistical Data on Wealth and Income', *A.E.R.*, VII (1917), supplement, Part I, 157–71; discussions by D. R. Dewey, p. 172; R. L. Hale, pp. 174, 175; and G. E. Roberts, p. 174

'Income and Wealth', *A.E.R.*, XV (1925), 457–74; discussion of questions relating to definitions and measurements

'Wealth Distribution on the Continental United States at the Close of 1921', *J.A.S.A.*, XXII (1927), 135–53

M. O. Lorenz, 'Methods of Measuring the Concentration of Wealth', *P.A.S.A.*, IX (1905), 209–19: discussions by G. K. Holmes, p. 318, and G. P. Watkins, p. 349

National City Bank, 'The Alleged Concentration of Wealth', *Monthly Letter*, Oct. 1936, pp. 147–51: wide distribution of products of industry cited as evidence of wide diffusion of wealth

'Who Owns the Wealth?', *Monthly Letter*, July 1936, pp. 103–7: again the consumption thesis

E. Z. Palmer, 'Sources and Distribution of Income in the South', *Southern Economic Journal*, Jan. 1936, pp. 47–60

W. M. Persons, 'Measurement of the Concentration of Wealth', *Q.J.E.*, XXIV (1909), 180–90

'The Variability in the Distribution of Wealth and Income', *Q.J.E.*, XXIII (1909), 416–49

LeGrand Powers, 'Factors Producing Unequal Wealth Distribution and Ownership and Measurements of Their Influence', *J.A.S.A.*, XXIII (1928), 417–28: the consequences of certain assumptions respecting wealth accumulation are traced

Hans Staehle, 'Short-Period Variations in the Distribution of Income', *Review of Economic Statistics*, XIX (1937), 133–43: an ingenious analysis using German statistics

Edgar Sydenstricker and W. I. King, 'Classification of the Population According to Incomes', *J.P.E.*, XIX (1921), 571–94

'The Measurement of the Relative Economic Status of Families', *P.A.S.A.*, XVII (1921), 842–57

Edgar Sydenstricker and G. St. J. Perrott, 'Changes in Family Income and Rental during the Economic Depression', *J.A.S.A.*, XXIX (1934), supplement, 42–6

W. Tresckow, 'Trust Business Possibilities: The Distribution of the Wealth of the United States and Potential Trustors', *Burrough's Clearing House*, Sept. 1931, pp. 13–15, 43, 44

R. S. Tucker, 'The Distribution of Income among Income Taxpayers in the United States, 1863–1935', *Q.J.E.*, LII (1938), 547–87

G. P. Watkins, 'The Growth of Large Fortunes', *Publications of the American Economic Association*, 3d series, VIII, No. 4 (1907), 735–904

'Measurement of the Concentration of Wealth', *Q.J.E.*, XXIV (1909), 160–79

Review of King's *Wealth and Income*, *A.E.R.*, VI (1916), 441–3

'Comment on the Methods of Measuring Concentration of Wealth', *P.A.S.A.*, IX (1904–05), 349–54

Edward White, 'Income Fluctuations of a Selected Group of Personal Returns', *J.A.S.A.*, XVIII (1922), 67–81

S. N. Whitney, 'Weakness of Data Supporting Conclusion of Increase in Diffusion of Wealth', *Annalist*, March 6, 1936, pp. 368, 369, 392 'Statistical Bases for National Wealth Estimates', *Annalist*, April 10, 1936, pp. 542, 562 (in letter form)

D. B. Yntema, 'Measures of Inequality', *J.A.S.A.*, XXVIII (1933), 423–33; summary of his Ph.D. thesis

A. A. Young, Review of King's *Wealth and Income*, *Q.J.E.*, XXX (1916), 575–87

Government Publications

Bureau of the Census, *Compendium of Eleventh Census: 1890* (Washington: Department of Commerce, 1897)

Bureau of Internal Revenue, *Statistics of Income* (Washington: Treasury Department, annually beginning in 1916)

L. J. Chawner, *Residential Building*, Housing Monograph Series, No. 1 (Washington: National Resources Committee, 1939)

J. A. Collins, *Distribution of Wealth in the United States*, Senate Doc. 75, 55th Cong., 2d Sess. (Washington, 1898)

Department of Agriculture, 'Returns Obtained from American Farmers from their Operations in 1922', *Weather, Crops and Markets*, April 28, 1923; a basic sample

Federal Home Loan Board, 'Incomes of Representative Home-Owner and Tenant Families in 61 Cities in 1933', *Review*, I (1935), 359–63

Federal Trade Commission, *National Wealth and Income*, Senate Doc. 126, 69th Cong., 1st Sess. (Washington, 1926)

G. K. Holmes and J. S. Lord, *Farms and Homes: Proprietorship and Indebtedness in the United States at the Eleventh Census, 1890* (Washington, 1896)

'Investigations of Bureau of Internal Revenue', *Senate Report No. 27*, 69th Cong., 1st Sess., Part II

Massachusetts Bureau of Statistics of Labor, *Twenty-fifth Annual Report* (1894)

National Resources Committee, *Consumer Incomes in the United States* (Washington, 1938)

Edgar Sydenstricker *et al.*, 'The Income Cycle in the Life of the Wage Earner', *Public Health Reports*, reprint No. 947, Aug. 22, 1924

Edgar Sydenstricker and W. I. King, 'A Method of Classifying Families According to Incomes in Studies of Disease Prevalence' in *Public Health Reports*, reprint No. 623, Nov. 26, 1920, pp. 2829–46

D. L. Wickens, *Financial Survey of Urban Housing* (Washington: Department of Commerce, 1937)

Bibliographies

There is no published bibliography with the precise scope required by the subject of this paper. However, helpful bibliographies are contained in several references cited above, for example: Berridge, Winslow, and Flynn, *Purchasing Power of the Consumer;* Leven, Moulton, and Warburton, *America's Capacity to Consume;* Streightoff, *Distribution of Incomes;* Yntema, *Measures of the Inequality;* and Watkins, 'Growth of Large Fortunes'. In addition, a bibliography on 'Income' compiled by M. T. Olcott of the Bureau of Agricultural Economics (Agricultural Economics Bibliography No. 73, May 1938) contains selected references on the concept of income and methods of obtaining income statistics.

Discussion

I SIMON KUZNETS

Mr. Merwin's survey reveals the variety of purposes that motivated the construction of distributions of income and wealth by size; and describes exhaustively the daring feats of ingenuity performed by skillful statisticians in their attempts to overcome the absence of basic information. In view of the lack of basic data, it seems surprising not that the estimates have been so few, but that there were any at all; not that they were so poor, but that they came within hailing distance of the truth.

However, the matter that interests me most is not the character and quality of the estimates, but the factors that explain the absence of basic information in the field. Why was no information collected during these decades on a sufficiently comprehensive scale to make possible an acceptable distribution of income or wealth by size among individuals or families? A great deal of other basic economic information was being collected, largely by public agencies: the censuses of population, agriculture, manufacturing; reports by the Interstate Commerce Commission on most public utilities; by banking authorities on the state of the credit system; by custom house authorities on foreign trade; and the like. Since, after all, the economic system functions in order to satisfy the needs of the nation's ultimate consumers, is it not surprising that information on what the economic system produced was not supplemented by equally important data on the flow of incomes to individuals or families, or on the stock of wealth at their command?

Mr. Merwin suggests briefly some of the factors that may serve to account for this gap in the economic information in the past and, to a less extent, even at present. But this point needs further discussion and illumination. While the supply of basic economic

data may be affected partly by accidental events, it is on the whole determined by fundamental views of the body social as to the relative importance of various aspects of economic activity and the need of information to aid in the solution of various economic problems. Data collection is expensive, both in the narrow sense of costing money to the collecting and the reporting agencies and in the broader sense of effort needed to translate the frequently unformulated and unandited impressions or records into reportable and quantitatively measurable facts. If in this country during recent decades, public agencies have been collecting so many data on some aspects of economic activity and so few on others, there must have been good and sufficient reasons. It is important to ascertain these reasons, for they provide clues to the factors that determine the supply of statistical data—a problem close to the heart of every empirically minded student in the social sciences. We shall, therefore, proceed to a necessarily tentative consideration of these reasons, with particular reference to the data on distribution of income by size among individuals or families.

It may be suggested that the path of progress in the collection of statistical data in the economic field is from population to production, and from production to distribution. It seems natural that the collection of census data by any nation would begin with the quantitative aspects of population, of people as the substance of the nation and the *ultima ratio* of its existence; would then proceed to ascertain what these people produce; and would concern itself with the distribution of results of economic activity among individuals or families only after having ascertained how many of them there are and how much they produce. One could thus say that the basic reason for the absence during recent decades of comprehensive information on distribution of income by size, concurrent with an apparently plentiful supply of data on production, is that generally the former would be collected later than the latter; and that this country was still in the phase of statistical development at which data on production could not yet be complemented by data on the distribution of the results of such production among ultimate consuming units.

Whether this stage-theory of the development of comprehensive statistical data is valid in terms of the actual historical ex-

perience in modern countries, I do not venture to say. One does get the impression that censuses of population and of production, in the order named, are the earlier phases in the growth of statistics in the various countries since the industrial revolution; but a careful test is beyond the scope of the present comments. In the absence of such careful tests, and of supplementary evidence, it would be impossible to demonstrate that with respect to this generally valid succession of stages, this country must have been in the second, the development of production statistics, and has failed as yet to reach the third, the development of data on distribution of results of production among consuming units. Nor would there be much meaning in such a statement.

But whether or not the generalization is historically valid, it should still be indicated why the development of various bodies of data should be sequential rather than concurrent. Why should data on distribution of income by size wait until data on population and production are complete, rather than be developed concurrently with the latter? The answer to this question seems to be that with scarce means, some selection of fields of comprehensive coverage must be made; that knowledge of one aspect of activity is an indispensable prerequisite for planning the statistical coverage of another; that the concurrent collection of several bodies of data is not necessarily complementary in terms of reducing per unit costs but may, on the contrary, serve to raise such costs; and that many statistical data are byproducts of the administrative activity of the government and hence are necessarily selective since governments cannot deal directly with everything at once. That in this necessarily sequential relation, data on population and production should precede those on distribution and consumption seems plausible.

This general impression of the primacy of production data and of the study of production processes has perhaps been reenforced by the rapid industrial development of this country during the decades under consideration. This resulted in a rather widely entertained, and, to a considerable extent justified, notion that the country's economic progress, i.e., increase in total product and economic power, was rapid; that the potentialities of such progress in the future were still considerable; that whatever problems might exist in the distribution of the national in-

come among the consuming units of the nation would be solved by the rapid rise of the production curve; and that correspondingly the function of the government was to preserve that freedom of private enterprise which would allow it to continue its splendid contribution to social welfare by raising the state of technical arts, extending the area of economic activity, and increasing the total of goods produced. Such an attitude meant that data collected for the purpose of information and observation would relate primarily to production—as a basis of judging the rate of progress and its origin in the various industries. It also meant that the public agencies were to be concerned primarily with the preservation of free competition within the country and maintenance of preferred position against foreign competitors —functions that involved dealing primarily with production agencies and hence collecting primarily production statistics. And data on production, being available largely within and hence provided almost exclusively by producing or business units, cannot yield data on distribution of income by size among consuming units.

If this was the viewpoint of society at large, there was also little pressure for income information on the part of business groups. Problems of marketing and distribution had not yet come to occupy the center of attention that they seem to now: the rapid extension of the productive system and growth in the volume of output meant that the restriction of the markets was a sporadic rather than a chronic circumstance. The growth in quality goods and semi-monopolistic markets, of advertising pressures and marketing surveys, was still largely in the future. The relation between income levels and consumption was of less importance to the business community at large than it is now; and thus one of the effective forces now pressing for information upon distribution of income by size, combined with regional breakdowns, was lacking, or at least much weaker than it is now.

The attitude of the individual to the provision of information on income was to a large extent a corollary of the general view of the body social on the greater importance of increasing production (and population) than of remedying the ills of income distribution, both products of the free individualistic organization of economic activity. Naturally enough, the people who

were at the top of the income pyramid resisted attempts to shed too much light on the inequalities in the distribution of income; and they continue to do so. But their resistance could not and cannot be successful unless it is backed by a negative attitude to the revelation of income information, an attitude that is a direct corollary of a viewpoint suggested above.

The connection is not difficult to see. If one believes that the economic system is enjoying and will enjoy in the future a rapid growth of output that will overcome any transient ills resulting from inequality in the distribution of income; if one believes that this beneficent progress is due to the invisible hand of providence which converts the selfish striving of individuals to their economic aggrandisement into a horn of plenty for the country at large; and if one considers further that part of this selfish behavior of free individuals is to withhold information of any sort, unless required by the state in order to perform its proper functions—then the reluctance to supply income information can be fully understood. The state should not do anything about income distribution, since the recipient of large income is being rewarded for his greater contribution to the national product and the recipient of small income is being punished for his failure to contribute. Since the state does not require income data of this type for the prosecution of its administrative activity; and since the unequal distribution of income is just a tool, and an efficient one, in stimulating economic growth, there is no reason why the free individual should sacrifice his competitive right to withhold information. The man who thought or was forced by society to think that he was the captain of his economic destiny would naturally resist giving an account for it to anyone but to his economic soul.

Technical obstacles undoubtedly added to the difficulties. To begin with, comprehensive coverage of any information relating to individuals or family units in the economy is much more difficult than coverage of productive or business units, for the simple reason that there are so many more of the former. Second, and perhaps more important, it is far more difficult to obtain accurate quantitative information from a consuming unit than from a member of the business system, since the accounting of the former is much more sketchy.

In this connection, it should be noted that of the various types of income, the one on which it is most difficult to obtain accurate information is that of individual entrepreneurs. In the case of salaries and wages, dividends or interest, the overt receipt of a payment makes it possible to recognize income clearly and to ascertain its magnitude with relative ease. The establishment of net income of individual entrepreneurs is a heroic task indeed. In the decades when individual entrepreneurs bulked so large among the income-earning population of the country and when even corporations were often unaware of the exact magnitude of their net income, it would have been difficult to survey family incomes in the same way as one establishes in the Census the age or sex of individual members of the families.

There is another, admittedly conjectural, consideration of great bearing upon the present status and prospects of the field: a distribution of income by size among families, for a single year or only a few years, and without many corollary data, is of limited value in the analysis of either policy or economic problems. Such a statement may seem at first surprising. However, brief reflection will show that even though great human interest attaches to a distribution showing that in a given year there were *x* families, each having an income of a million dollars and over, and *y* families, each having an income of less than one thousand dollars, such an estimate taken by itself for a year or two is scarcely illuminating. Of course, such estimates are used, but ordinarily on the dangerous assumption that the distributions for one year hold for a longer period; that differences among various income classes in cost of living, size of families, or other factors are not significant for interpreting income differences for the purpose at hand—and there are very few purposes for which such an assumption is tenable; and that there is enough stability within the distribution from year to year to allow a rough identification of families within a given income category with the same families within a similar income category at another time.

Of course, it is questionable that the realization of the low value of a distribution for a single year, unaccompanied by many corollary data, was clearly in the minds of the people who were in a position to determine whether comprehensive data in the field would be collected. But it would not be unreasonable to

assume that this feeling, namely, that unless one could initiate a continuous series of such estimates and obtain both the necessary breakdowns and the subsidiary data the effort was not worth making, did serve to reduce the pressure and to prevent sporadic collections of data.

The striking additions to the data on distribution of income by size during very recent years tend to bear out the tentative analysis above of the factors that made for the absence of such data in the past. The accumulation of information on population and production and their intensive use by students in the field resulted in relatively satisfactory knowledge of these aspects of the economy. Of course, significant gaps remain even in these fields, especially on some of the dynamic elements: population migration, production of intangible goods, scale of producing unit and of business unit, etc. But a great deal of further work in these fields must await better data on distribution of income among consuming units, data the absence of which is felt perhaps more acutely than ever before by students whose major interest is not income measurement or in the analysis of closely related economic problems.[1]

More obvious is the change in social attitude and in the economic functions of government as they are now conceived by society at large. The feeling that there are great reserves of production growth in the future is not widely entertained now; and therefore, to put it mildly, serious doubts are entertained as to the future effectiveness of the system of free and individualistic economic organization. That this organization is largely a thing of the past, a result of the growth of private and semi-public monopolies, is a significant element in the changed situation. And there is less conviction that the economic fortune of an individual is entirely or even largely a result of his personal ability. It is realized that the complex of economic institutions does not function perfectly or even tolerably well, and that these imperfections have painful repercussions among large groups in our society, repercussions these groups could not cope with or avoid by any individual effort, no matter how well meant or intelli-

[1] It is important that the advocacy of income questions on the tentative population schedule for the 1940 Census came from population statisticians interested in the economic factor in differential fertility.

gently designed. Correspondingly, the economic fortunes of those favorably situated are seen as due only in part, and perhaps small part, to their personal ability to contribute to social welfare: they are seen as being to a large extent a result of strategic situations created by social institutions and seized upon by a few individuals, often to the detriment of society at large.

It is this viewpoint that provides the pressures and justification for activities of public authorities designed to modify the working of economic institutions in their determination of the distribution of income among individuals and families. It provides the *raison d'être* for a graduated income tax, social security legislation, laws concerning wages and hours, etc. It thus brings government into fields of administrative activity whose byproducts are large bodies of data on distribution of income by size. And it creates an attitude on the part of the community at large that makes the provision of income information a natural and acceptable step designed to help the public authorities in dealing with a commonly recognized economic problem.

Furthermore, the increasing attention the business system at large pays to methods of gauging and influencing the consumers' market results in pressures, often effective, for information on distribution of income by size and on related expenditures. True, this particular drive is somewhat biased toward higher income brackets (as is true of the byproducts of income tax laws) and often leads to a somewhat exaggerated estimate of income magnitudes; but it is a potent factor, nevertheless, in forcing this field upon the attention of public authorities.

The technical difficulties in the way of collection of income information of the type under discussion are also becoming less formidable, partly because of the increasing importance of income in the form of overt payments, partly because the technical means at our disposal for dealing with large populations have increased at an undoubtedly greater rate than the population itself. Such means include not only the technical and organizational machinery for dealing with large scale surveys, but also the statistical theory that makes it possible to establish in advance the reliability of samples and thus to select the latter on a carefully thought out basis.

It seems quite probable that we are on the verge of a period

during which comprehensive data on the distribution of income by size and hence reliable distributions based on such data will become available. We may, therefore, be confronted in the very near future with choices among various ways in which such information can be obtained and distributions based on them derived. While deep-seated factors determine the feasibility or impracticability of obtaining comprehensive statistical information on this or another phase of social activity, once these factors are favorable to the collection of such information, the academic student, guided by general interests only, is in a position to shape many of the evolving data and assure their greater usefulness in the treatment of the problems with which he is concerned. And this he can do by participating in the selection of the various alternatives that exist, either overtly or implicitly, when the task of comprehensive coverage of a field like distribution of income by size is initiated.

In this choice the consideration already stated, that distributions for single years, unaccompanied by many related data (on occupation, industry, family composition, age, sex, location, cost of living, expenditures and savings, etc.) are of little use, seems to me paramount. In the various choices two criteria should be given the heaviest weight, next to that concerning the basic reliability of the information likely to be obtained: the likelihood that the data will (1) become available continuously, on an annual basis or on the basis of relatively short time units; (2) be obtained in such a way that correlation with many other factors will be possible. We may be disillusioned by the low analytical value of the first distributions, since their great significance will become obvious only after they have cumulated into a long series and have been tested for association with variables other than income. But unless we assure that such development will be possible, our disillusionment is likely to become permanent.