

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

Volume Title: Shares of Upper Income Groups in Income and Savings

Volume Author/Editor: Simon Kuznets, assisted by Elizabeth Jenks

Volume Publisher: NBER

Volume ISBN: 0-87014-054-X

Volume URL: <http://www.nber.org/books/kuzn53-1>

Publication Date: 1953

Chapter Title: The Basic Variant

Chapter Author: Simon Kuznets, Elizabeth Jenks

Chapter URL: <http://www.nber.org/chapters/c3064>

Chapter pages in book: (p. 279 - 300)

Chapter 8

THE BASIC VARIANT

In Chapter 7 we discussed the difficulty of comparing income on tax returns with total income receipts of individuals because the former excludes some income items that should be included and includes some that should be excluded. The major difficulty, however, lies in the way the income tax data are classified.

The tabulations published annually in *Statistics of Income* and available in more detail in the *Source Book* for the years beginning with 1927, classify tax returns by their net income, defined in most years as total income, including capital gains reportable by law, minus allowable deductions (tax and interest payments, capital losses, etc.), but not reduced either by personal exemptions or credits for dependents (or by prior year loss, even when permitted as an offset in calculating the tax).¹

For our analysis such classification suffers from two major defects: the use of the return as a unit and of net income as a base. Since we are interested in income per person rather than per return, we need a classification by income per person. And since an individual's share in countrywide income payments is measured properly by his income receipts excluding such transfer items as capital gains, and not reduced by any transfer losses or by deductions that may be interpreted as part of the cost of living (i.e., nonbusiness expenses), and including such imputed income as net rent on owner-occupied dwellings, the income total that should be used in the classification by size is substantially different from net income, tax definition.

We must either discard the available classification, except for the kind of comparison in Chapter 7, or adjust it as best we can to fit our needs. We adopted the second course, and the remaining chapters of Part IV discuss our procedures, devices, and assumptions. This chapter deals with

¹The income total underlying the classification of returns by size changes from one part of the period to the next as the tax law and form of return change. In recent years, e.g., returns on the brief 1040A form are tabulated by size of gross income, i.e., total income gross of any deduction, and from 1944 on all returns are classified by total income minus allowable trade and business deductions. But by and large, for most years the basis of classification by size is net income, which includes capital gains and is reduced appreciably by deductions that are a legitimate part of economic income.

the estimates of shares of upper income groups as derived from the available material with only such modifications as still permit us to use it in almost complete detail. The shares are designated basic because they permit fullest analysis and hence are those most emphasized in Part I. The adjustments applied to this basic variant in an attempt to reach more closely the desired approximation to shares in economic income are discussed in Chapters 9 and 10.

1 *Basic Variant, Total Population*

The basic variant yielding the shares of upper income groups of the total population is derived by the following procedure. For each net income class, tax definition, in the published data we calculate, by methods already discussed, economic income on returns — income which, with minor qualifications,² accords with that underlying countrywide income payments. Likewise, for each net income class we estimate the total number of persons covered by the returns. We can, therefore, calculate per capita economic income for each net income class (with the qualifications listed in note 2). Next we array the classes in decreasing order of economic income per capita, cumulate their population and income from the top down, then express each succeeding level of these distributions as a percentage of countrywide population and income respectively. In these cumulative distributions we interpolate at the points where the tax return population constitutes 1, 3, 5, 7, and for some years, larger percentages of total population. Since these interpolations are made in the cumulative distributions, counting from the top, the line drawn at 1 percent of total population cuts the cumulated income distribution at a point that shows the percentage of countrywide income payments received by the 1 percent of the population drawing the highest incomes. The 3, 5, 7, etc. percent lines of total population yield corresponding shares of income.

A sample calculation of the share of the top 1 percent in this basic variant for total population is given in Appendix 3, Section A. Several technical points may be noted here. First, the 1 percent of population line is in most years well below the \$10,000 net income class, tax definition, and no class above this line can affect our estimate as the array is reshuffled when the classification base shifts from net income to economic income. Hence we start with returns with net income of \$10,000 and over,

² These qualifications are: (a) before 1939 wages and salaries of nonfederal employees are omitted; (b) imputed net rent on owner-occupied dwellings is omitted throughout; (c) income derived through financial intermediaries without current distribution, i.e., insurance companies and the like, is omitted; (d) the reincluded deductions contain some business expenses that should not be reincluded in precise calculations. Tentative adjustments for items (a) and (b) are presented and discussed in Chapter 9.

treating them as a single class; and distinguish and check upon the order of only the net income classes below \$10,000.³ Second, the income totals for the successive income classes are rearranged before cumulation, then converted into percentages of the countrywide total. Third, the interpolation is by a straight line to the logarithms, i.e., direct proportions of logarithms of percentages of population and of income. This is justified by the fact that the two cumulative percentage series plotted on a double logarithmic chart form straight lines for almost all ranges and practically each year in the period.

From these interpolations at successively lower percentages we derive the share of income received by the top 1 percent, the next pair, the 2nd and 3rd percentage band, and so on, until we are stopped by the limits of the data. The chief aspects of the procedure, of possible interest to the technically minded reader, are presented in Table 75.

As may be noted in columns 1 and 2, in many years the lowest partition line, i.e., the lowest percentage of population line in the cumulative distribution from top incomes down, is drawn appreciably short of the total tax return coverage. This is a safety measure since evasion and under-reporting may be more widespread at the exemption limits and filing margins than at some distance above them. And the series used in the analysis in Part I stop short, for all years except 1930, 1931, and 1933, even of the partition lines in column 2. On the other hand, the desire to have a comparable continuous series for as long a period as possible has tempted us to include 1913-16, when the total coverage is barely 1 percent of population and in 1915 even less; and to draw the lowest partition line in other years, notably 1917, 1918, 1927-29, and 1934-35, uncomfortably near the very bottom of the tax return distribution. In general, the shares of the lower percentage bands in all these years are subject to a wider margin of error than those for years when we stop appreciably short of exhausting the full tax return population.

Column 3 indicates the extent of shifting in the array produced by ranking net income classes, tax definition, by economic income per capita instead of by net income, tax definition, per return. Two conclusions emerge. First, changes in rank are few, confined in most years to one that affects only two income classes. This is obviously due to the fact that the class intervals distinguished in the data are wide: in most years for \$1,000 spans. Only for 1941 through 1943, when the spans in the lower ranges

³ This economy of labor is not justified when the top 1 percent line is close to or above the \$10,000 and over class limit of net income, tax definition, or were we to distinguish partition lines at higher ranges, e.g., the top 0.25 percent. When the former occurs, we extend our analysis to cover the necessary additional net income classes; the latter refinement is difficult with the available data, as is explained below.

Table 75

Coverage of the Basic Variant, Total Population, 1913-1948

	% of Total Population Covered by Tax Return Population (1)	Lowest % Line Covered by Variant (2)	Classes Changing Rank in Shift from Net Income, Tax Defini- tion, Per Return to Economic Income Per Capita (thousands of dollars) (3)	Net Income Class, Tax Defini- tion, in Which Lowest % Line Lies (dollars) (4)	Economic Income Per Capita, Lowest % Band Distin- guished (5)	Economic Income Per Capita, Tax Return Population Omitted (6)
1913	1.05	1	"	"	\$5,013	\$1,114
1914	1.04	1	"	"	4,206	953
1915	0.96	1	"	"	4,852	...
1916	1.22	1	None	3-4	6,209	1,307
1917	7.38	7	1-2 2-3	2-3	955	863
1918	10.44	10	1-2 2-3	2-3	840	759
1919	11.83	10	1-2 2-3	2-3	966	854
1920	16.28	10	1-2 2-3	2-3	1,016	814
1921	13.59	10	Under 1 2-3 3-4	2-3	1,043	948
1922	14.27	10	Under 1 2-3	2-3	987	901
1923	16.60	10	1-2 2-3	2-3	1,052	943
1924	15.04	10	1-2 2-3	2-3	1,045	1,009
1925	8.14	7	2-3 3-4	1-2	1,366	1,173
1926	8.04	7	2-3 3-4	1-2	1,371	1,208
1927	7.61	7	2-3 3-4 4-5	1-2	1,401	1,325
1928	7.50	7	2-3 3-4	1-2	1,406	1,135
1929	7.61	7	2-3 3-4	1-2	1,445	1,275
1930	6.83	5	2-3 3-4	3-4	1,509	1,229
1931	5.96	5	2-3 3-4	1-2	1,410	1,071
1932	7.09	7	None	Under 1	923	683
1933	6.86	5	None	1-2	990	820
1934	7.25	7	None	Under 1	961	695
1935	7.78	7	None	1-2	1,056	859
1936	8.93	7	None	1-2	1,095	984
1937	10.19	10	None	Under 1	1,115	783
1938	9.93	7	None 1-2	1-2	1,107	1,079
1939	11.73	10	2-2.5 2.5-3 3-4	2.5-3	1,120	1,056

Table 75 concluded:

	% of Total Population Covered by Tax Return Population (1)	Lowest % Line Covered by Variant (2)	Classes Changing Rank in Shift from Net Income, Tax Definition, Per Return to Economic Income Per Capita (thousands of dollars) (3)	Net Income Class, Tax Definition, in Which Lowest % Line Lies (4)	Economic Income Per Capita, Lowest % Band Distinguished (5)	Economic Income Per Capita, Tax Return Population Omitted (6)
1940	24.03	20	1-2 2-2.5 2.5-3	2-2.5	\$907	\$809
1941	43.81	20	8 Classes ^b	2.5-3, Form 1040A	912	759
1942	62.84	20	17 Classes ^b	2-2.25, Forms 1040 & 1041	952	729
1943	71.40	20	11 Classes ^b	2.5-2.75, Forms 1040 & 1041	1,119	792
1944	80.36	20	None	3-3.5	1,152	766
1945	82.90	20	None	3-3.5	1,125	734
1946	86.88	20	None	3-3.5	1,179	765
1947	88.25	20	None	3.5-4	1,280	865
1948	87.59	20	None	4-5	1,394	917

^a Data not available by income classes.

^b The classes are as follows (thousands of dollars):

1941		1942		1943	
Form 1040A	Forms 1040 & 1041	Form 1040A	Forms 1040 & 1041	Form 1040A	Forms 1040 & 1041
0.75-1	0.75-1	0.75-1	0.75-1	1.5 -1.75	1.25-1.5
1 -1.5	1 -1.5	1 -1.25	1 -1.25	2 -2.25	1.5 -1.75
1.5 -2	1.5 -2	1.25-1.5	1.25-1.5	2.25-2.5	1.75-2
2 -2.5		1.5 -1.75	1.5 -1.75	2.5 -2.75	2 -2.25
2.5 -3		1.75-2	1.75-2	2.75-3	2.25-2.5
		2 -2.25	2 -2.25		2.5 -2.75
		2.25-2.5	2.25-2.5		
		2.5 -2.75	2.5 -2.75		
		2.75-3			

Column

- 1 Table 69, column 6.
- 2 From Table 113, column 2.
- 3 For 1929 see columns 1, 6, and 8 of the sample calculation in Appendix 3, Section A. Entries for the other years are derived by the same procedure.
- 4 The total income received by the lowest percentage band is the product of total income receipts of individuals (Table 72, col. 2) and its share in that total (Table 118, col. 1). Its total income is then divided by its population (for total population, see Table 69, col. 5) to yield its income per capita.
- 5 From economic income covered by tax returns (Table 72, col. 1) we subtract the amount received by all the upper percentage bands distinguished (estimated by the procedure indicated in the notes to col. 5 for the amount received by the lowest percentage band). From the population covered by tax returns (Table 69, col. 3) we subtract the population covered by all the upper percentage bands (the product of total population, Table 69, col. 5, and col. 2 of this table). The income not covered by these percentage bands is then divided by the population not covered.

are reduced to \$250 and when the published data separate short from long forms (1040A on the one hand, and 1040 and 1041 on the other), does the number of income classes changing rank increase markedly. In 1944 and later years, when all returns are classified by adjusted gross income, approximating our concept of economic income, no shifts occur.

The second conclusion is that the changes in rank are among low income classes: in years when they occur they are, with a single exception, for classes below \$4,000 net income, tax definition. This indicates that of the two factors that can produce a shift in rank — difference between net income, tax definition, and economic income as bases of classification, and difference between returns and persons as units of classification — the latter is by far the more important. We observed in Chapter 7 that the difference between the two income bases was relatively more appreciable in the upper net income brackets, tax definition, than in the lower ones, except the very bottom; hence this factor would tend to produce shifts in rank chiefly among the upper net income brackets. But because most single person returns are in the lower income brackets, there may be appreciable differences from bracket to bracket in the number of persons per return, and it is this factor that is chiefly responsible for the shifts in rank at the lower income levels. A check of this conclusion by the detailed information in Tables 111-113 confirms its validity.⁴

The per capita economic income of the lowest percentage band distinguished in the basic variant (col. 5) is larger than that of the residual part of the tax return population (col. 6). This is a necessary arithmetical result of the arraying procedure used; what is of interest is the relative difference between the two. In general, it runs well above 10 percent, being especially large in the early years when the top 1 percent is the only band covered; in the early 1930's, when there is a rapid falling away of income

⁴ This check involves identifying in Table 113, column 2, the net income classes that changed rank; then observing for them columns 2 and 8 in Table 111, and the columns in Table 112 that show net income, tax definition, and economic income. Thus, for 1917 we note in Table 113 that the net income class \$1,000-2,000 shifted above the \$2,000-3,000 class. In Table 111, we find that for the \$1,000-2,000 class, the tax return population is estimated to be 2.3 million for 1.6 million returns; for the \$2,000-3,000 class, 2.4 million for only 0.8 million returns. Thus for the former class, the number of persons per return is 1.4; for the latter class, 2.8. In Table 112, net income, tax definition; for the \$1,000-2,000 class is \$2.46 billion and economic income, \$2.66 billion — an increase of about 8 percent; the corresponding totals for the \$2,000-3,000 class are \$2.06 and \$2.24 billion respectively, an increase also of about 8 percent. Clearly, the change in rank is due to the number of persons per return factor, not to the shift in income base. Every shift in rank can be similarly traced and analyzed from Tables 111 through 113.

below the last band included; and in 1941-48, when the lowest percentage line falls far short of the total coverage of the tax return population.

One aspect of the procedure, common to the basic variant and to all adjustments of it, is not revealed by Table 75: the limiting of percentage bands to 1, 3, 5, 7, 10, etc. and the avoiding of more narrowly defined bands either above or below the top 1 percent line. The reason is the roughness of the classification of the published data which for most years distinguish only \$1,000 intervals of net income, tax definition, in the lower ranges of the distribution. When the class intervals are few and broad, the distinction of narrower percentage bands in our analysis would mean several interpolations within one and the same net income class; and any differences among shares of successive percentage bands might be spurious. This argument does not apply to refinement *within* the top 1 percent band, since the published distributions give much detail for class intervals above its lower partition line; however, as capital gains and various deductions become more important, estimating a proper distribution by economic income becomes much more difficult than in the lower brackets, and the resulting approximations subject to a much wider margin of error. It was, therefore, considered best not to push the analysis into partition lines above the top 1 percent of total population.⁵

2 Coverage of Farm Population by Tax Returns

In Chapter 7 we compared the tax return population and its economic income with total population and countrywide income payments. The basic variant just discussed is also in terms of shares of the upper percentage groups of total population. Might it not be just as relevant, at least in the way of supplementary analysis, to compare the number and income of the tax return population with the nonfarm population and its income?

This question is strongly suggested by even a superficial study of the tax returns, especially their classification by the residence of the filer and by the industrial source of the income reported. Some evidence that the returns are, through most of the period, overwhelmingly from nonfarm areas was given in Chapter 7. Another measure of the extent to which the farm population is covered is the number of returns reporting income from farming. An attempt to estimate its proportion in all net income returns is provided in Table 76.

Several assumptions had to be made to fill in gaps in the evidence; of these, three are vital for understanding the results. First, we assumed that the income from farming entered on returns that reported it constituted

⁵ However, the basic variant for nonfarm population does involve shares of a top percentage that is less than 1 percent of total population (Sec. 3).

Table 76

286 Number and Income of Returns Showing Income from Farming, and Number and Income of All Tax Returns, 1918-39, 1941
(includes only net income returns; returns and population in thousands; dollar figures in millions)

	Returns Showing Net Profit from Business										
	All Net Income Returns (1)	All Business Industries (2)	Agr. & Related Industries (3)	% Is of Col. 1 (4)	Population Represented by Col. 1 (5)	Col. 3 (6)	% Is of Col. 5 (7)	Economic Income, All Returns (8)	Net Profit from Business All Business Industries (9)	Agr. & Related Industries (10)	% Is of Col. 8 (11)
1918	4,425	958	372	8.41	10,922	1,200	10.99	17,591	3,124	1,123	6.38
1919	5,333	1,124	419	7.86	12,441	1,326	10.66	21,647	3,878	1,211	5.60
1920	7,260	964	272	3.74	17,361	902	5.19	25,931	3,216	637	2.46
1921	6,662	807	130	1.96	14,776	398	2.69	23,162	2,366	229	0.99
1922	6,787	906	105	1.54	15,726	336	2.14	24,141	2,840	231	0.96
1923	7,698	1,746	185	2.41	18,610	590	3.17	28,323	4,723	357	1.26
1924	7,370	1,646	165	2.23	17,179	526	3.06	28,274	4,755	372	1.32
1925	4,171	979	83	1.99	9,433	265	2.81	22,545	3,689	250	1.11
1926	4,138	944	79	1.91	9,443	257	2.72	23,292	3,573	239	1.02
1927	4,102	859	71	1.73	9,070	219	2.41	23,533	3,287	237	1.01
1928	4,071	774	63	1.54	9,046	197	2.18	24,407	3,244	234	0.96
1929	4,044	839	67	1.66	9,266	218	2.35	25,392	3,328	216	0.85
1930	3,708	750	45	1.20	8,411	140	1.67	21,178	2,628	117	0.55
1931	3,226	628	20	0.61	7,389	61	0.83	16,869	1,890	38	0.22
1932	3,877	585	17	0.44	8,852	54	0.62	14,325	1,295	24	0.17
1933	3,724	626	33	0.89	8,613	105	1.22	12,950	1,403	54	0.42
1934	4,094	655	50	1.23	9,159	154	1.68	15,044	1,717	97	0.65
1935	4,575	660	56	1.22	9,896	168	1.70	16,954	1,855	124	0.73
1936	5,413	751	68	1.26	11,435	203	1.77	21,039	2,374	169	0.80
1937	6,350	788	71	1.12	13,126	208	1.59	24,146	2,493	175	0.73
1938	6,204	839	82	1.32	12,889	236	1.83	21,631	2,349	171	0.79
1939	7,633	921	97	1.27	15,351	276	1.80	26,072	2,711	202	0.77
1941	25,855	2,970	1,089	4.21	58,351	3,205	5.49	64,029	6,453	1,767	2.76

Notes to Table 76

Column

- 1 Table 67, column 1.
- 2 1918-25, 1937, 1939, and 1941: *Statistics of Income*. Data for 1939 and 1941 include returns with no net income and exclude taxable fiduciary returns.
1928: number as reported on Form 1040 in *Statistics of Income* plus the number estimated for Form 1040A by dividing net profit from business for each income class (*Statistics of Income*) by the mean income for the given income class, estimated as the arithmetic average of its upper and lower levels.
1926, 1927, 1929-36, and 1938: number as reported in *Statistics of Income* for returns with net incomes \$5,000 and over plus the number estimated for returns with net incomes under \$5,000, calculated as follows:
Net profit from business for each income class under \$5,000 (*Statistics of Income* or the *Source Book*) is divided by the mean income for the given class, estimated as the arithmetic average of its upper and lower levels. For 1926, 1927, and 1938 the number of returns with net profit from business thus derived is regarded as final. For 1929-36 it is adjusted by the average of the 1928 and 1937 ratios of the actual number of such returns with net incomes under \$5,000 (*Statistics of Income*) to the number just computed.
- 3 1918-25, 1939, and 1941: *Statistics of Income*; see note to column 2 regarding 1939 and 1941.
1928: number as reported on Form 1040 in *Statistics of Income* plus the number estimated for Form 1040A on the assumption that the industrial distribution of all business returns on Form 1040A (see notes to col. 2) is the same as that of business returns on Form 1040 with net incomes under \$5,000 (*Statistics of Income*).
1926, 1927, and 1929-38: estimated by the following steps:
 - 1) For 1925, 1928, and 1939 the percentage that column 3 constitutes of column 2 is computed.
 - 2) For 1925-37 *Statistics of Income* shows the industrial distribution of returns with net incomes of \$5,000 and over reporting net profit from business. The percentage that the number reporting net profit from agriculture and related industries constitutes of the total number is computed.
 - 3) For 1926 and 1927 the percentage derived in step 1 is interpolated between 1925 and 1928 along a straight line. For 1929-37 it is extrapolated from 1928 by the series derived in step 2. For 1938 it is interpolated along a straight line between the percentage for 1937, just derived, and the percentage for 1939 calculated in step 1.
 - 4) Column 2 is multiplied by the percentages derived in step 3.
- 5 Table 69, column 3.
- 6 Column 3 multiplied by the average number of persons per family return, Table 68, column 1.
- 8 Table 72, column 1.
- 9 *Statistics of Income*; see note to column 2 regarding 1939 and 1941.
- 10 1918-25, 1939, and 1941: *Statistics of Income*; see note to column 2 regarding 1939 and 1941.
1928: Net profit as reported for returns on Form 1040 in *Statistics of Income* plus net profit for returns on Form 1040A estimated as the product of the number of such returns (see notes to col. 3) and the net profit per return, calculated by the following steps:

(continued on page 288)

their total economic income; second, that the units filing these returns were members of the farm population; third, that these returns were family returns. The first assumption leads to an underestimate of the income of the farm population reported on tax returns. The second leads to an overestimate which would offset, if only in part, the underestimate from the first assumption. The last assumption is the least doubtful of the three in that farming is a family business, particularly at the income levels that would entail filing a federal income tax return.

If the assumptions are valid, and the error implied in them cannot be fatal, we can accept the estimates in Table 76 as representing the total number and income of the farm population recorded on tax returns. The proportion the farm population constitutes of the total represented on tax returns is quite small, somewhat over 10 percent at the highest and ranging from 1 to 3 percent in most years. The proportion of income attributable to the farm tax return population is even lower, not exceeding 7

Notes to Table 76, column 10, concluded:

- 1) Net profit from agriculture and related industries per return is calculated for returns on Form 1040 with net income under \$5,000.
 - 2) Net business profit per return is calculated for all returns on Form 1040 with net incomes under \$5,000.
 - 3) The ratio of the net profit per return derived in step 1 to that derived in step 2 is calculated.
 - 4) Net business profit per return is calculated for all returns reporting such profits on Form 1040A (see notes to col. 2).
 - 5) Net business profit per return as calculated in step 4 is multiplied by the ratio derived in step 3.
- 1926, 1927, 1929-38: estimated in 9 steps:
- 1) For 1925-39 net profit from business per return reporting it is calculated by dividing column 9 by column 2.
 - 2) For 1925, 1928, and 1939 net profit from agriculture and related industries per return reporting it is calculated by dividing column 10 by column 3.
 - 3) For 1925-37 net profit from business per return reporting it with net incomes \$5,000 and over is calculated from *Statistics of Income*.
 - 4) For 1925-37 net profit from agriculture and related industries per return reporting it with net incomes of \$5,000 and over is calculated from *Statistics of Income*.
 - 5) For 1925, 1928, and 1939 the ratio of the net profit per return as calculated in step 2 to that calculated in step 1 is computed.
 - 6) For 1925-37 the ratio of the net profit per return calculated in step 4 to that calculated in step 3 is computed.
 - 7) For 1926 and 1927 the ratio derived in step 5 is interpolated between 1925 and 1928 on the basis of the change in the ratio derived in step 6. For 1929-37 it is extrapolated from 1928 with the ratio in step 6 as index. For 1938 it is interpolated along a straight line between the ratio for 1937, just calculated, and that for 1939 derived in step 5.
 - 8) Net business profit per return calculated in step 1 is multiplied by the ratio derived in step 7.
 - 9) The number of returns in column 3 is multiplied by the net profit per return calculated in step 8.

percent at the highest and below 1 percent in most years. As the coverage of tax returns widened after 1941, the percentages accounted for by the farm population and its income must have increased appreciably; but even in these years the proportions they constitute of the upper income groups (confined, say, to the top 5 percent of total population and its income) are, on the basis of the record, within the low levels indicated for most years by columns 7 and 11.

One conclusion is obvious: tax returns may be treated as filed almost exclusively by the nonfarm population, especially if we emphasize the upper percentage bands and discard, in any calculated variants, the lower tail of the tax distribution. In other words, we can legitimately compare population and economic income on tax returns not only with total population and its income but with the nonfarm population and its income, thereby deriving a variant that yields the shares, year in year out, of the upper percentages of the nonfarm population. These estimates will be subject to error as far as some members of the farm population are included whose omission might have led to replacement by members of the nonfarm population with perhaps different per capita income. But the error is obviously slight enough so that the reliability of the basic variant for the nonfarm population is not appreciably less than that for the total population.⁶

Before discussing this basic variant for nonfarm population, we must touch upon a different question raised by Table 76: do the low proportions of farm tax return population and its income in the total tax return population and its income reflect genuine differences in income levels between farm and total population, or are they due to more evasion and underreporting by the farm population? If the former, the biases in the basic variant for total population due to underreporting and evasion are relatively equal to those in the basic variant for nonfarm population. If the latter, those in the basic variant for total population are greater than those in the basic variant for nonfarm population.

A valid answer could be given only if we had size distributions of economic income per person, separately for the farm and nonfarm populations, both distinguishing fairly narrow size classes, especially in the upper

⁶ The additional error involved in the basic variant for nonfarm population is not measured by the percentages in columns 7 and 11 of Table 76. If we could exclude the returns that report income from farming, they would be replaced by returns from additional members of the nonfarm population (to fill out to the proper percentage of the nonfarm population whose share is being estimated). The estimate would, therefore, be modified only because the nonfarm units shifted into the given percentage band might have a slightly smaller per capita economic income than that on the returns from the farm population excluded. The implicit error would thus be only a minor fraction of the percentages in Table 76.

ranges. We could array the size classes, cumulate them in a single distribution from the top down (keeping the farm and nonfarm distinction for subsequent recognition), then interpolate at the top 1, 3, etc. percentage lines. We could then ascertain what proportion of the farm population should characterize the top x percent, x being the proportion of the total population accounted for by the total tax return population for the given year. Comparison of the proportion of the farm population in the top x percent of total population, thus calculated, with the proportion of the farm tax return population in the total tax return population for that year (given in Table 76, col. 7), would tell us whether the nonreporting bias for the farm population was bigger than for the nonfarm.

Such data are unavailable, even the NRC estimates for 1935-36 not permitting a rearranging of the income distributions for the farm and nonfarm populations by income per person. But an experimental test, starting with some bold assumptions, was attempted. The basic assumption was that the inequality in the size distribution of income was the same for the farm and nonfarm populations — inequality being measured by the ratio of the average income of the given income group to the average income of the population, as revealed for each year by its share in the basic variant for total population. Thus, if the share of the top 1 percent of the total population is 13 percent, the ratio of its average income to the average income of the entire population is as 13 to 1, and the ratio of the average income of the top 1 percent of the farm (or nonfarm) population to the average income of the entire farm (or nonfarm) population is also as 13 to 1.

With the help of this and some auxiliary assumptions we constructed two hypothetical distributions, one for the nonfarm and the other for the farm population, and arrayed them in such a way as to see how large a proportion of the upper income groups is accounted for by the latter. An illustrative calculation for 1929 is given in Appendix 3, Section B. All we need to note here is that this calculation is likely to exaggerate the proportion of the farm population that should be represented on the tax returns — for two reasons. First, the relative excess of the average income of upper percentage groups over the average income of the total population is likely to be larger for the nonfarm than for the farm population merely because the spread of income opportunities is much wider for the former and hence the probability of very large incomes is much greater.⁷ Second, the

⁷ This, however, might be more than offset by the effect of the preponderance of entrepreneurial income in the income of the farm population, which may make for greater dispersion and inequality in the size distribution of income of the farm population (see Ch. 6, Sec. 3).

procedure by which the shares of the lower percentages of the nonfarm population are estimated is likely to place their income levels too low in the combined array.

The results are erratic from year to year, and only the average for the entire period is of interest. If the assumptions are correct, the proportion of the farm population that should be reported on tax returns averages about 6 percent for the period. Since our entries in Table 76 average only 3 percent, they suggest that there is relatively more underreporting among farm income recipients than among nonfarm.

But large as the difference seems, the effect on the reliability of the basic variant for total population as compared with that for nonfarm is not marked. All that the experimental calculation shows is that the basic variant for total population is subject to a somewhat larger bias of underreporting — a rough maximum on the average of 3 percent of the total coverage. In other words, the relatively greater underreporting by the farm group affects only 3 percent of the total tax return population which might have reported somewhat larger incomes than those actually reported. As already remarked, the error in estimating upper group shares resulting from such displacement can be only a minor fraction of the percentage that gauges the displacement itself.

While the average level of the proportions the farm tax return population and its income constitute of the total tax return population and its income (Table 76, col. 7 and 11) is probably too low, the changes in these proportions are confirmed by independent data. In Table 77 and Chart 5 are estimates showing the relation between the income of the entire farm population and total income receipts. These countrywide estimates are completely independent of data on federal tax returns from individuals whereas all the percentages in Table 76 are derived from them. Hence, when we compare the percentage that the income of the farm population constitutes of the income of the total population with the percentage the farm tax return population constitutes of the total tax return population or with the percentage the income on farm tax returns constitutes of economic income reported on all tax returns, we are comparing independent series.

It is this fact that makes the close agreement in Chart 5 so significant. In lines a, b, and c the percentages the farm tax return population and its income constitute of the total tax return population and its income move in close conformity with the percentage the income of the farm population constitutes of the income of the total population. Likewise, in lines d and e the shortage of the proportion of income on farm tax returns relative to the proportion of population represented changes in close conformity with

Table 77

Percentages that Population and Income on Farm Returns Constitute of Total Tax Return Population and Income Compared with Percentage that Income of Farm Population Constitutes of Total Income Receipts, 1918-39, 1941

	Income of Farm Population as % of Total Income Receipts (1)	Per Capita Income (dollars)		Ratio: Col. 2 to Col. 3 (4)	Ratio of % of Income to % of Population, Farm to Total Tax Returns (5)
		Farm Population (2)	Nonfarm Population (3)		
1918	19.51	339	610	0.56	0.58
1919	18.46	377	693	0.54	0.53
1919	18.29	377	702	0.54	0.53
1920	14.86	314	759	0.41	0.47
1921	11.36	191	614	0.31	0.37
1922	11.51	208	646	0.32	0.45
1923	11.37	243	727	0.33	0.40
1924	11.75	255	708	0.36	0.43
1925	12.52	287	727	0.39	0.39
1926	11.58	279	750	0.37	0.38
1927	11.43	281	737	0.38	0.42
1928	11.14	280	747	0.37	0.44
1929	11.11	295	778	0.38	0.36
1930	9.72	231	697	0.33	0.33
1931	8.61	168	581	0.29	0.27
1932	8.00	118	450	0.26	0.27
1933	9.81	137	432	0.32	0.34
1934	11.05	179	487	0.37	0.39
1935	11.78	209	521	0.40	0.43
1936	11.52	234	584	0.40	0.45
1937	11.37	255	628	0.41	0.46
1938	10.97	229	573	0.40	0.43
1938	10.90	229	577	0.40	0.43
1939	10.54	238	613	0.39	0.43
1941	11.29	345	787	0.44	0.50

Column

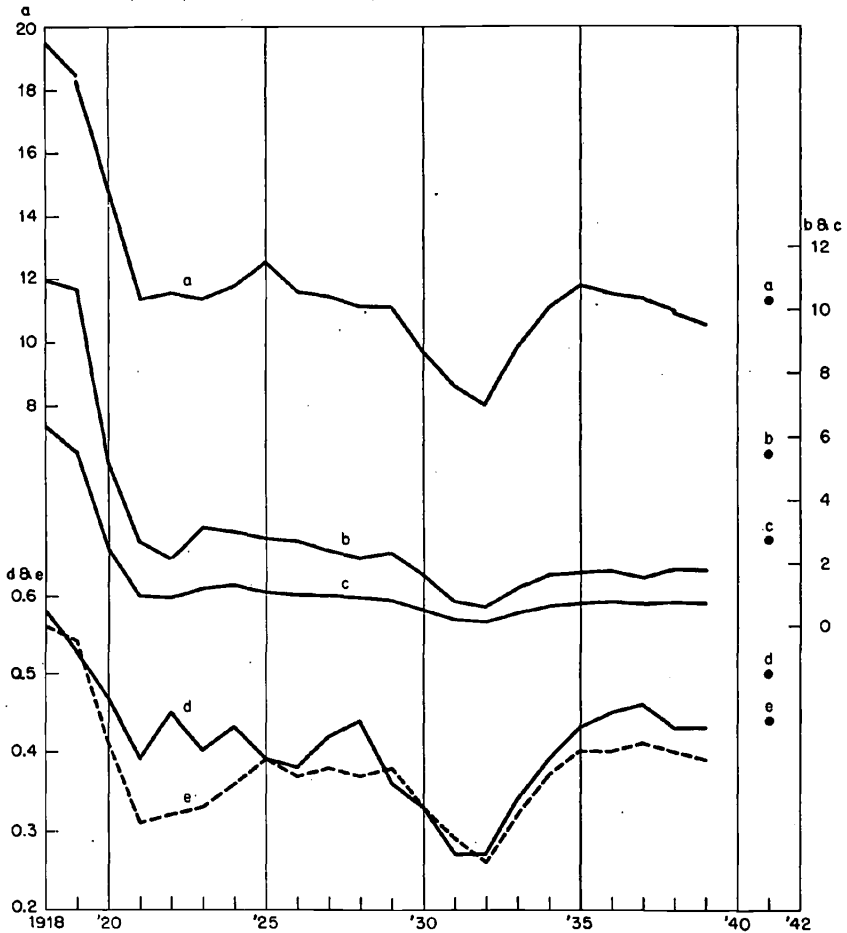
- 1 Income of farm population (Table 114: difference between column 12 and column 13) divided by total income receipts of individuals (column 12).
- 2 Income of farm population (see note to column 1) divided by farm population (difference between column 5 of Table 69 and column 1 of Table 115).
- 3 Table 115: column 2 divided by column 1.
- 5 Table 76: column 11 divided by column 7.

the shortage of the per capita income of the farm population relative to the per capita income of the nonfarm population. In other words, our estimates for the tax return population reflect faithfully both the changing proportion of the income of the farm population in the income of the total population and the changing inequality in average per capita income between the farm and nonfarm population. If there are any differences in

Chart 5

Percentages that Population and Income on Farm Returns Constitute of Total Tax Return Population and Income Compared with Percentage that Income of Farm Population Constitutes of Total Income Receipts 1918-39 and 1941

- a Income of farm population as % of total income receipts
- b Farm tax return population as % of total tax return population
- c Income of farm tax return population as % of income of total tax return population
- d Ratio: c to b
- e Ratio: per capita income of farm population to that of nonfarm population



the reporting bias between the farm and nonfarm population, they obviously do not vary enough to conceal genuine changes in the distribution of total income between the two groups or in the inequality between their per capita incomes.

3 *Basic Variant, Nonfarm Population*

As we have seen, the proportion of the farm tax return population in the total tax return population is quite small, and its proportion in the upper percentage bands would be even smaller. It certainly seems clear that the tax returns covered in our estimate of the top 1 percent of the total population include so few, if any, farm returns that it is justifiable to treat all as being from the nonfarm population, and hence to compare them with the total number and income of the latter. This interpretation of tax returns as being almost exclusively from the nonfarm population would probably be quite justifiably applied also to those included in the 2nd and 3rd and in the 4th and 5th percentage bands of the basic variant, and, with continuously decreasing confidence, to those in the lower percentage bands.

It is impossible to say at what point in the distribution the proportion of farm tax returns in all tax returns becomes large enough to render this interpretation untenable. In the detailed tables in Part V this point was taken to lie at the 10th percentage line from the top in the basic variant for nonfarm population. In some years this point is probably too low; and for purposes of analysis in Part I we stopped at the 7th percentage line. Within the tax return population above this line the proportion of farm tax returns is assuredly small enough to permit us to treat all as being for the nonfarm population alone.

We compute the basic variant for nonfarm population by a procedure analogous to that employed in deriving the basic variant for the total population except that we now use the number and income of the nonfarm population as denominators. In the array of classes by economic income per capita, derived from the tax return tabulations, we interpolate lines at 1, 3, 5, etc. percent from the top of the nonfarm population,⁸ cutting off the percentages of income received by its ordinal percentage bands, and yielding the income shares of the top 1, 3, 5, etc. percentage bands. An illustrative calculation for 1929 is provided in Appendix 3, Section C.

What the basic variant for nonfarm population shows concerning the level of and changes in the inequality of the size distribution of income was

⁸ The interpolation using straight lines on a double logarithmic scale follows the procedure used for the basic variant for the total population. But in 1938 the slope of the straight line changed materially from one interclass interval to the next, and some smoothing was called for to yield successively decreasing percentages of income per unit as we went down the array. The smoothing was done simply by omitting some of the interclass intervals, thereby drawing the straight line over a wider interval than in the fully detailed array. Because the smoothing was applied to the nonfarm variant alone, the results of the comparison of the basic variant for the total and the nonfarm population for 1938 are not strictly in line with the results for other years.

discussed in Part I. Here we are concerned with the technical aspects of the procedure that explain how and why the differences between income shares in the basic variant for the total and the nonfarm population come out as they do. As noted in Part I, the difference between income shares of identical percentage bands in the basic variant for the total and the nonfarm population was relatively narrow; and while the shares of the top 1, and 2nd and 3rd percentage bands were, on the whole, larger in the variant for the nonfarm population, the opposite was, on the whole, true of the share of the 4th and 5th percentage band (and also, in the shorter period covered, of the 6th and 7th percentage band).

The average per capita income of the nonfarm population is larger than that of the total population. If, for a given group of tax returns in the upper brackets we compare economic income per person with the average income of a wider population group, the excess of the former would be relatively less if the latter were the average income of the nonfarm population than if it were the average income of the total population. Therefore, in the shift from the basic variant for the total population to that for the nonfarm, the percentage share of income of an identical group of returns, disregarding the proportion of population represented, would be lower in the nonfarm variant. The ratio of the per capita income of the nonfarm population to the per capita income of the total population (Table 78, col. 1), thus measures a factor that would make the share of a given percentage band in the basic variant for nonfarm population smaller than its share in that for the total population.

But we must take into account also the proportion of the population represented. If the nonfarm population is only 60 percent of the total population, a group of returns that is included in the top 1 percent of the former is, at the same time, a group that covers only the top 0.6 percent of the latter. This top 0.6 percent of total population must, in the nature of the case, have a higher per capita income than the group of returns that comprises the full top 1 percent of the total population. The ratio of the per capita income of the top 0.6 percent of the total population (equal to the top 1 percent of the nonfarm population) to that of the top 1 percent of the total population measures the factor that would make the share of a given percentage band (in this case, the top 1 percent) in the nonfarm variant larger than its share in the basic variant for the total population. This raising factor, which differs in size for the several percentage bands, is shown for the top 1, 2nd and 3rd, 4th and 5th, and combined top 5 percent, in Table 78, columns 2-5.⁹

⁹ An illustrative calculation of the raising factor for 1929 is provided in Appendix 3, Section D.

Table 78
Analysis of Relative Difference between Shares of Upper Income Groups in Basic Variant for Nonfarm Population and Those for Total Population, 1913-1948

	Ratio: Per Capita Income of Nonfarm to That of Total Population					Ratio: Per Capita Income of Given Percentage Band of Nonfarm Population to That of Total Population					Ratio: Income Share of Given Percentage Band of Nonfarm Population to That of Total Population						
	(1)	(2)	(3)	(4)	(5)	Top 1	2nd & 3rd	4th & 5th	Top 5	Top 1	2nd & 3rd	4th & 5th	Top 5	Top 1	2nd & 3rd	4th & 5th	Top 5
1913	1.27	1.37								1.08				1.08			
1914	1.26	1.36								1.08				1.08			
1915	1.25	1.35								1.08				1.08			
1916	1.24	1.35								1.09				1.09			
1917	1.20	1.31	1.21	1.07	1.24					1.10	1.01	0.89	1.04	1.01	0.89	1.04	1.07
1918	1.16	1.28	1.21	1.13	1.23					1.11	1.04	0.98	1.07	1.01	0.98	1.07	1.07
1919	1.16	1.25	1.25	1.17	1.24					1.08	1.08	1.01	1.07	1.01	1.01	1.07	1.07
1919	1.16	1.25	1.25	1.17	1.24					1.08	1.08	1.01	1.07	1.01	1.01	1.07	1.07
1920	1.21	1.26	1.24	1.18	1.24					1.04	1.02	0.97	1.02	1.02	0.97	1.02	1.02
1921	1.25	1.25	1.23	1.08	1.21					1.00	0.98	0.86	0.97	1.00	0.86	0.97	0.97
1922	1.24	1.25	1.19	1.20	1.22					1.00	0.96	0.97	1.00	1.00	0.96	0.97	0.98
1923	1.23	1.24	1.20	1.09	1.20					1.01	0.98	0.89	0.98	1.01	0.98	0.89	0.98
1924	1.21	1.22	1.19	1.13	1.20					1.01	0.99	0.94	0.99	1.01	0.99	0.94	0.99
1925	1.19	1.21	1.21	1.11	1.19					1.02	1.02	0.93	1.00	1.02	0.93	1.00	1.00
1926	1.20	1.21	1.24	1.07	1.19					1.01	1.04	0.89	1.00	1.01	1.04	0.89	1.00
1927	1.19	1.21	1.22	1.04	1.18					1.02	1.03	0.88	1.00	1.02	1.03	0.88	1.00

1928	1.18	1.20	1.21	1.06	1.18	1.02	1.02	0.90	1.00
1929	1.18	1.20	1.20	1.08	1.18	1.02	1.02	0.91	1.00
1930	1.20	1.20	1.20	1.03	1.16	1.00	1.00	0.86	0.97
1931	1.21	1.19	1.15	1.06	1.15	0.98	0.95	0.87	0.95
1932	1.22	1.20	1.12	1.07	1.15	0.98	0.92	0.88	0.94
1933	1.21	1.20	1.10	1.13	1.16	0.99	0.91	0.94	0.96
1934	1.19	1.20	1.17	1.07	1.16	1.00	0.98	0.90	0.97
1935	1.18	1.20	1.18	1.04	1.16	1.02	1.00	0.88	0.98
1936	1.17	1.20	1.19	1.03	1.16	1.02	1.02	0.88	0.99
1937	1.17	1.19	1.18	1.08	1.17	1.02	1.01	0.93	1.00
1938	1.17	1.18	1.16	1.21	1.18	1.01	1.00	1.04	1.01
1938	1.16	1.18	1.16	1.21	1.18	1.01	1.00	1.04	1.01
1939	1.17	1.17	1.18	1.01	1.14	1.01	1.01	0.87	0.98
1940	1.16	1.17	1.16	1.16	1.17	1.01	1.00	1.00	1.00
1941	1.14	1.16	1.16	1.14	1.16	1.02	1.02	1.00	1.01
1942	1.11	1.16	1.17	1.12	1.15	1.04	1.05	1.00	1.04
1943	1.10	1.14	1.15	1.12	1.14	1.04	1.05	1.02	1.04
1944	1.09	1.12	1.15	1.11	1.13	1.03	1.06	1.02	1.04
1945	1.09	1.11	1.14	1.12	1.12	1.02	1.05	1.03	1.04
1946	1.09	1.12	1.15	1.13	1.13	1.03	1.05	1.04	1.04
1947	1.09	1.12	1.14	1.13	1.13	1.03	1.04	1.03	1.03
1948	1.08	1.12	1.13	1.12	1.12	1.03	1.04	1.03	1.04

Column

1 Column 3 of Table 77 divided by column 5 of Table 74.

Column

2-5 For 1929 see sample calculation in Appendix 3, Section D.
6-9 Table 116: column 4 divided by column 1.

Chart 6

Relative Difference between Income Shares of Upper Income Groups in Basic Variant for Nonfarm Population and those for Total Population, and Factors Determining Its Magnitude, 1913-1948

- a Ratio: income share of given percentage band of nonfarm population to that of total population
- b Factor a: ratio of per capita income of nonfarm population to that of total population
- c Factor b: ratio of per capita income of given percentage band of nonfarm population to that of total population

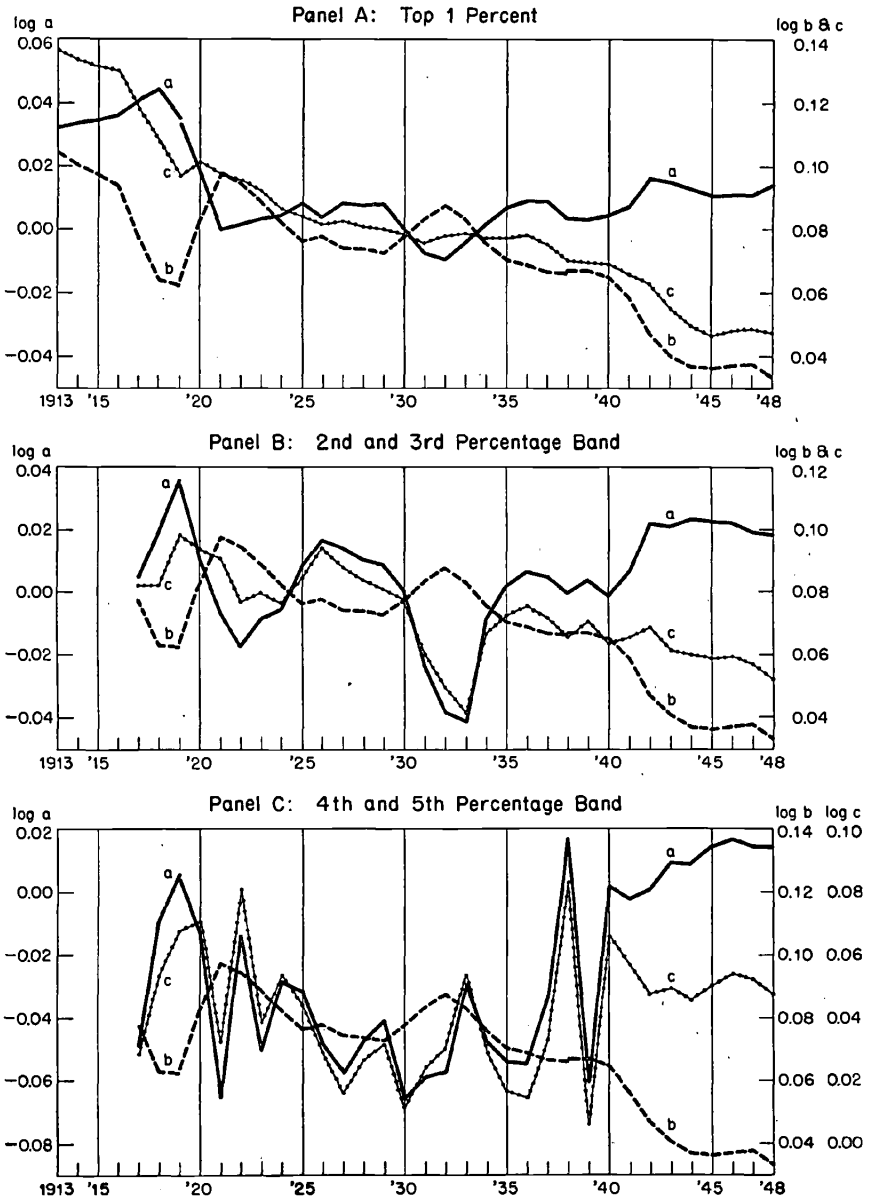
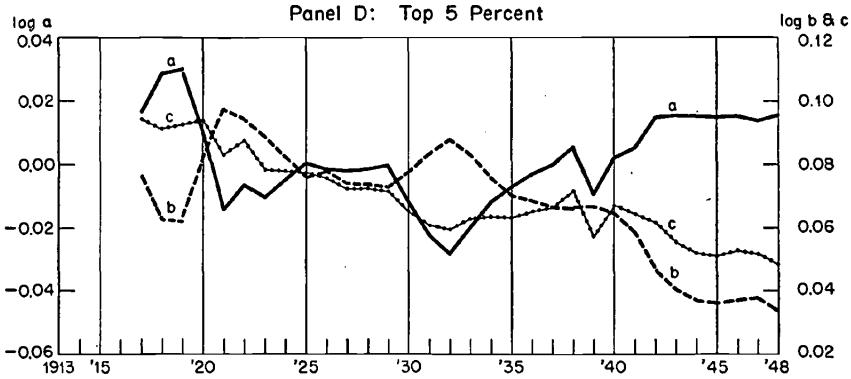


Chart 6 (concl.)

- a Ratio: income share of given percentage band of nonfarm population to that of total population
- b Factor a: ratio of per capita income of nonfarm population to that of total population
- c Factor b: ratio of per capita income of given percentage band of nonfarm population to that of total population



The relative difference between the income shares of identical upper percentage bands in the basic variant for the nonfarm and the total population is then a compound of two factors: the ratio of the per capita income of the nonfarm to the per capita income of the total population, which tends to make the shares in the variant for the former smaller, and the ratio of the per capita income of the percentage band in the total population that is the equivalent of the given upper, x , percentage band in the nonfarm population to the per capita income of the upper x percentage band in the total population, which would always raise the share of an upper percentage band of the nonfarm population above the share of the ordinarily identical percentage band of the total population. Dividing the second (the raising) ratio by the first (the reducing) ratio should yield the ratio, for ordinarily identical percentage bands, of the share in the nonfarm variant to the share in the variant for the total population. For example, in Table 78, the entry for each year in column 6 is the result of dividing the ratio in column 2 by that in column 1; and it is at the same time identical with the ratio of the share of the top 1 percent in the nonfarm variant to that in the variant for the total population (see entries for these shares in Table 116, col. 4 and 1 respectively).

With the help of Table 78 and Chart 6 we can see why the share of a given percentage band in the variant for the nonfarm population differs from that in the variant for the total population. First, the difference is relatively small obviously because the reducing and the raising ratio tend largely to offset each other. As the share in the nonfarm variant is

reduced because the denominator base — the average income with which the upper incomes are being compared — is larger than in the variant for the total population, this reduction is offset (or more than offset) because the numerator — the per capita income of the upper group — is also raised by drawing the partition lines at ordinal percentages that segregate much narrower upper groups in the total population.

Second, the relative size of the reducing and the raising ratio differs for the successive percentage bands in the nonfarm variant. For the top 1 percent, and somewhat less for the 2nd and 3rd percentage band, the raising ratio exceeds the reducing ratio. Hence, the share of the top 1, and somewhat less, of the 2nd and 3rd percentage band of nonfarm population are, on the whole, larger than the shares for the identical percentage bands in the variant for the total population. For the 4th and 5th percentage band, and also for the 6th and 7th (not shown in the table or chart) in the nonfarm variant, the reducing ratio is, on the whole, larger than the raising ratio. Hence, the shares of these percentage bands in the nonfarm variant are, on the whole, smaller than their shares in the variant for the total population.

Third, both ratios decline because the percentage of the nonfarm population in the total population has been increasing fairly steadily, except in a few years in the depressed 1930's, when some people went back to the farm. Obviously, if the nonfarm population approaches the total population in size, its per capita income too, other conditions being equal, will tend to approach that of the latter; and an ordinal percentage band in the former will tend to approach an ordinaly identical percentage band in the latter. Hence both ratios, as far as they always tend to be above 1 (with 1 as the lower limit), will show a downward drift as the proportion of the nonfarm population in the total rises.

Consequently, the ratio of the share in the nonfarm variant to its share in the variant for total population, for an ordinaly identical percentage band, will also approach unity, i.e., the relative difference will tend to disappear as the nonfarm population approaches the total population in size. On the whole then, this ratio should be nearer 1 in the later years of the period studied than in the earlier. This trend is confirmed by Table 78 and Chart 6 until the years associated with World War II. The rise in the ratio in the recent years is probably due to an increase in the proportion of farm returns in the upper brackets, i.e., to the failure of our basic assumption. Such an increase in the proportion of farm returns, influencing the numerator of the share but not its denominator, would tend to raise the ratios in columns 6-9 above 1.