

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

Volume Title: Studies in Income and Wealth

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

Volume Publisher: NBER

Volume ISBN: 0-870-14168-6

Volume URL: <http://www.nber.org/books/unkn51-1>

Publication Date: 1951

Chapter Title: Field Surveys of Consumer Income: An Appraisal

Chapter Author: Robert Wasson, Abner Hurwitz, Irving Schweiger

Chapter URL: <http://www.nber.org/chapters/c5730>

Chapter pages in book: (p. 482 - 559)

# Part IX

## Field Surveys of

---

*The authors gratefully acknowledge their indebtedness to the staffs of the Bureau of the Census, Bureau of Labor Statistics, Federal Reserve Board, and the Survey Research Center of the University of Michigan for valuable criticisms, suggestions, and unpublished data; also to the members of the Interdepartmental Technical Committee on Income Distribution, who gave generously of their time in assisting in the preparation of this paper. Any errors of fact or interpretation are, of course, the responsibility of the authors.*

*The Federal Reserve Board field surveys were conducted for the Board of Governors by the Survey Research Center of the University of Michigan. The staff of the Center was responsible for the detailed planning and supervising of the survey and the interview. At the time of the first survey many of the present Center staff were associated with the Department of Agriculture, Division of Program Surveys.*

# Consumer Income An Appraisal

*Robert Wasson*

NATIONAL INCOME DIVISION, DEPARTMENT OF COMMERCE

*Abner Hurwitz*

BUREAU OF LABOR STATISTICS

*Irving Schweiger*

BOARD OF GOVERNORS OF THE FEDERAL RESERVE SYSTEM

---

SINCE THE 1935-36 Study of Consumer Purchases, five federal agencies have collected data on consumer annual total money income. These data were obtained in eight nationwide surveys and cover five years. None of these studies was designed primarily for the collection of income data. In the 1941 survey of consumer income and expenditures the Bureau of Labor Statistics collected urban data and the Bureau of Human Nutrition and

Home Economics, rural data. The Census surveys of population, labor force, and housing for 1944, 1945, and 1947 covered farm and nonfarm units. In 1946 the Bureau of Agricultural Economics covered the farm population; the Census, urban and rural nonfarm. The Federal Reserve Board in its consumer finance studies for 1945, 1946, and 1947 was concerned chiefly with consumer savings and their expected use. In addition, the BLS collected urban income and expenditure data for 1944 and data in 3 cities for 1945, 1946, and 1947; the Census collected income statistics for Washington, D.C. for 1947; and the Bureau of Home Economics conducted several local area surveys of family income and expenditures.

The surveys were similar in several respects: the general collection procedure; definitions of income; and definitions of the income-receiving unit, which was either a group of related persons residing in the same dwelling unit or some subdivision, based upon arrangements for pooling finances or expenditures.

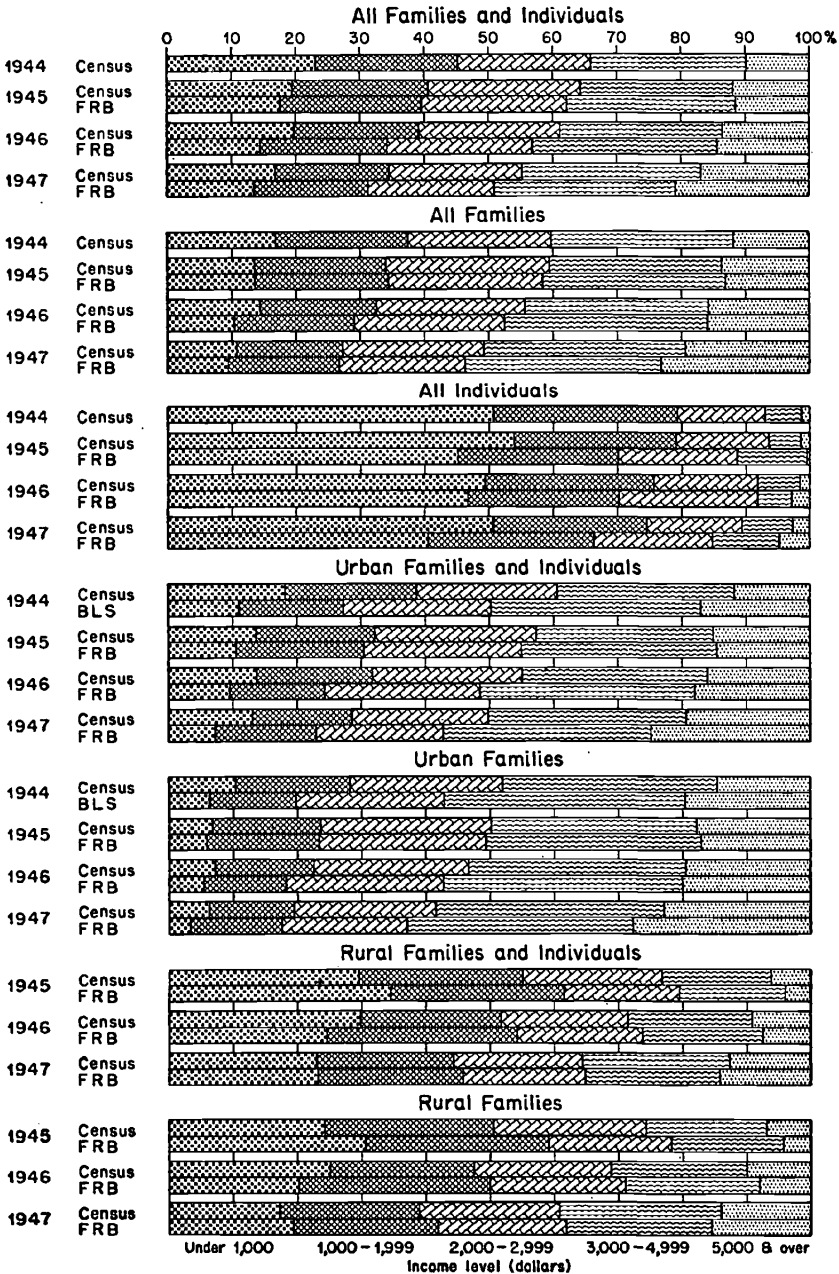
However, a comparison of published income figures derived from these surveys would be confusing to the uninitiated, because the definitions and the universes, though similar, were not identical. The definitions and terminology used even by the same agency have varied from one survey to another. In addition, the agencies have collected and processed the data by different techniques.

This paper describes the differences in definition, terminology, and techniques employed in the nationwide surveys for 1944-47 and the Washington, D.C. BLS and Census studies, and attempts to give these differences quantitative expression. Unfortunately, sufficient information is lacking with which to reconcile the results of the various surveys completely, but major technical differences in procedures are discussed.

#### A SUMMARY OF CONCEPTS AND PROCEDURES

All samples were area as distinguished from quota samples; households were selected systematically, not left to the interviewers' choice. The Census nationwide surveys of 1944-47

# All, Urban, and Rural Families and Individuals Percentage Distribution by Income Level



covered 6,700–12,000 households; the FRB nationwide surveys of 1945–47 covered 2,600–3,200 families and individuals; and the BLS 1944 urban survey covered about 1,700 families and individuals.

A dwelling unit was a room or group of rooms meeting various standards, such as separate kitchen facilities, use as separate living quarters by one or more persons. Units not meeting the standard and in structures containing at least 10 similar units were termed rooming units.

A household consisted of all persons occupying a dwelling unit. Transient hotels, large rooming houses, etc., were termed quasi-households.

The family, as defined by the Census and FRB, consisted of two or more persons residing in the same living quarters and related by blood, marriage, or adoption. The treatment of deceased persons and children away at college differed somewhat. Individuals or individuals not in families were persons living alone or with persons not related to themselves. The BLS economic family consisted of 2 or more persons not necessarily related who lived together sharing their expenses and dependent upon a common or pooled income, or a 1-person unit called a 1-person family or single consumer.

The BLS attempted to reconstruct the family as it existed during the survey year. The Census and FRB accepted the family composition as of the date of interview, with one minor exception.

Only money income was covered. Imputed income and capital gains and losses were excluded. All commonly accepted sources of money income were included by all agencies, but the treatment of rent and farm income differed somewhat. The detail for nonfarm units varied considerably. For 1944–46 the Census used from 7 to 11 questions, dropping to 2 in 1947. The FRB used 5 questions in 1945 and 12 questions in 1946 and 1947. The BLS schedule was always detailed.

The Census used a record-type of question for 1944–46 and a modified interview form for the nationwide and Washington 1947 surveys. Except in 1947 it left the exact formulation of the

income questions to the interviewers. The FRB printed open-end interview-type questions as they were to be asked by interviewers. The BLS used interview-type questions, leaving the exact form to the interviewers.

Each agency checked its returns for completeness and consistency; when necessary, units were reinterviewed for missing data or to eliminate inconsistencies. The most thorough check for consistency was the provision on the BLS schedules for detailed information concerning income, expenditures, and savings.

In the 1944 and 1945 Census surveys and all FRB surveys the weighted sample results were inflated to agree with independent estimates of the number of dwelling units in the survey universe. For 1946 and 1947 the Census inflated its weighted sample results to agree with independent estimates of the population of the United States with respect to certain demographic characteristics, including age and sex.

## B EVALUATION OF SURVEY RESULTS

### 1 *Census and BLS Data*

The 1944 urban and the 1947 Washington surveys afford the only BLS income data that can be compared directly with Census income figures (Table 1). Although each agency modified the design of the survey the distributions for 1944 and 1947 differed consistently. Both BLS distributions had considerably higher frequencies in the upper income levels than the Census. The BLS median income for all income-receiving units was 20 percent higher than the Census in 1944; 31 percent higher in 1947. The differences in the distributions were larger for individuals than for families; but even for the latter they seem to have been beyond the range of sampling variability.

Lacking sufficient information with which to reconcile these differences, we attempted only to determine factors that might help explain them and the probable effect of reconciliation. The factors considered are survey coverage, definition of the income-receiving unit, family reconstruction, income data obtained, the memory factor, sample design, and survey data adjustments.

Table 1  
Families and Individuals, Percentage Distribution by Income Classes  
Census and BLS, 1944 and 1947

	1944 Urban U.S.				1947 Washington, D.C.			
	Census	BLS	Difference		Census	BLS	Difference	
			Absolute	%			Absolute	%
<i>Families</i>								
Under \$1,000	10.4	6.3	-4.1	-39.4	3.5	0.4	-3.1	-88.6
1,000-1,999	17.9	13.5	-4.4	-24.6	8.6	5.5	-3.1	-36.0
2,000-2,999	23.7	23.1	-0.6	-2.5	16.1	13.9	-2.2	-13.7
3,000-3,999	21.9	22.2	+0.3	1.4	18.9	17.2	-1.7	-9.0
4,000-4,999	11.5	15.3	+3.8	33.0	15.9	15.0	-0.9	-5.7
5,000-9,999	12.3	19.6	+5.0	34.2	32.4	42.9	+10.5	32.4
10,000 & over	2.3				4.6	5.1	+0.5	10.9
Total	100.0	100.0			100.0	100.0		
Median, \$	2,918	3,320	+402	13.8	4,162	4,867	+705	16.9
Standard error, \$*	55	100	114		60	194	204	
<i>Individuals</i>								
Under \$1,000	45.7	35.9	-9.8	-21.4	17.9	12.0	-5.9	-33.5
1,000-1,999	30.7	30.9	+0.2	0.7	22.3	16.0	-6.3	-28.0
2,000-2,999	15.2	22.8	+7.6	50.0	34.2	24.0	-10.2	-29.3
3,000-3,999	5.0	5.4	+0.4	8.0	14.0	28.0	+14.0	100.8
4,000-4,999	1.9	1.9	0.0	0.0	5.7	10.0	+4.3	75.0
5,000 & over	1.5	3.1	+1.6	106.7	5.9	10.0	+4.1	69.4
Total	100.0	100.0			100.0	100.0		
Median, \$	1,115	1,435	+320	28.7	2,261	2,917	+656	29.0
Standard error, \$*	60	130	143		62	292	299	
<i>All income-receiving units</i>								
Under \$1,000	18.2	11.0	-7.2	-39.6	7.7	2.2	-5.5	-71.4
1,000-1,999	20.6	16.3	-4.3	-20.9	12.5	7.1	-5.4	-43.2
2,000-2,999	21.8	23.0	+1.2	5.5	21.2	15.5	-5.7	-26.9
3,000-3,999	18.1	19.5	+1.4	7.7	17.5	18.9	+1.4	8.0
4,000-4,999	9.5	13.2	+3.7	38.9	12.9	14.2	+1.3	10.1
5,000-9,999	9.9	17.0	+5.2	44.1	24.6	37.8	+13.2	53.7
10,000 & over	1.9				3.6	4.3	+0.7	19.4
Total	100.0	100.0			100.0	100.0		
Median, \$	2,480	2,987	+507	20.4	3,401	4,444	+1,043	30.7
Standard error, \$*	50	90	103		50	198	204	

As the text indicates, the BLS definitions of family and individuals are not entirely comparable with those of the Census. Comparability is greater for the 1944 than for the 1947 distributions. It is believed, however, that the net effect of the incomparabilities in definition is to minimize the difference.

\* Standard errors of the difference between medians, calculated by the formula  $\sigma_D = \sqrt{\sigma_{2mdn}^2 + \sigma_{2mdn}^2}$ , appear in the absolute difference columns.

Both BLS surveys included persons who resided in households, lodging houses, hotels, dormitories, YMCA's, fraternities, etc., nurses' homes, tourist camps, trailers, or were noninmate residents of nonmilitary institutions. The 1944 Census sample was selected from residents of households, lodging houses, and trailer camps. It did not include hotels, dormitories, YMCA's, nurses' homes, etc., or noninmate residents of nonmilitary institutions, where a large proportion of individuals might be expected to be found. Such individuals probably did not receive



higher incomes than other individuals. Consequently, their inclusion would probably lower the median of the Census distribution of all units by weighting the individual distribution more heavily, thereby widening the differences in survey results. The Census did include these units in the 1947 Washington sample. This, combined with the effect of BLS eligibility requirements, which exclude individuals who have families living elsewhere, may in small part explain the larger differences in the Washington studies.

In the 1944 studies the Census units of 2 or more related persons and the BLS economic family of 2 or more persons differed, aside from reconstruction, only in that the BLS occasionally included a group of unrelated persons who pooled incomes. Since such groups were few, this difference probably had little effect on the results. In the 1947 study the BLS definition of the economic family allowed a division of the Census family of related persons into 2 or more groups of persons or individuals, depending upon their financial arrangements. Combining these economic family units into Census family units would throw more units into the higher income classes, raise the BLS median income, and widen the differences in the distributions.

During 1944 and early 1945 persons entering the armed forces still far outnumbered returning veterans; their income, and the income of civilians who died during the period, were not recorded by the Census. A large number of single migrants continued to leave existing family groups to establish new units, including servicemen's wives who rented rooms near training camps. The Census recorded them as separate income-receiving units; the BLS, by reconstructing the family, included them as members of the families they left. To the degree that these factors were operating, Census data understated the median income of units as they existed during the survey year. Other changes in family composition tended probably to maintain some balance in the distributions or possibly to cancel some of the downward bias in the Census data. Departing servicemen's wives joined existing families and the tightening housing market

led to some doubling up. On the other hand, returning veterans reconstituted their families, and newly formed and undoubled families occupied new and reconverted dwellings as fast as they became available.

Conditions in the Washington area in 1947 were typical of a prosperous economy. When the housing market eased somewhat, changes in family composition resulting from undoubling, etc. tended to increase the number of income-receiving units. The Census failure to reconstruct the family increased the units in the lower income classes and biased the median income downward.

By requesting the amount of income from each job, and details on wages and salaries (see App. B), the BLS possibly obtained more complete and accurate reports than the Census. By questions on payroll deductions errors in reported gross income from wages and salaries were discovered and corrected. Almost all such corrections raised the income figures, for respondents generally knew their take-home pay but were not always sure of the amount deducted from their pay checks. Since deductions were large in both 1944 and 1947, the reporting and recording of take-home pay as wage and salary income before deductions would bias the income distribution downward. This factor was no doubt less serious in the Washington survey than in the 1944 urban study, since earners in Washington are preponderantly per annum salaried employees who presumably know their base pay scale. Median earnings of principal and supplementary earners in the Washington area were estimated to be \$3,183 and \$1,488 respectively by BLS; \$3,146 and \$1,478 by the Census.

Income data for 1944 were collected by the Census in May, and by the BLS in January-April 1945. The 1947 surveys ran in the field simultaneously. Whether the memory bias due to the lapse of time between the end of the year and the date of the interview was greater in the Census 1944 income data is impossible to say; the more complete filing of income tax reports by April may have benefited the Census data. However, it may reasonably be assumed that the BLS 1944 data were biased less

than the Census because of the poor memory of the informant for 3 reasons. First, the more detailed BLS income questions aided the informant to remember income from irregular sources and to understand what was meant by wages and salaries. Second, the BLS collected, along with income data, information on expenditures and changes in family net worth. Total expenditures plus changes in net worth were checked against total reported income. If these totals did not balance to a reasonable degree, the BLS agent revisited the unit and reviewed the schedule entries with the respondent; quite often, forgotten income was remembered. Third, the BLS agent spent an average of 6 hours with the family in completing a schedule for income, expenditures, and savings. The length of the interview allowed time for many items of income to be recalled that might otherwise have been unreported. Questioning of individual earners in the family in connection with their income and expenditures, especially for clothing, probably helped to bring to light the earnings of occasional workers. In the Washington studies the BLS covered a significantly higher percentage of multiple-earner families than the Census.

Washington Family Units, Percentage Distribution by Number of Earners

<i>Number of Earners</i>	<i>Census</i>	<i>BLS</i>
0	3.4	1.5
1	54.6	43.2
2	33.2	42.5
3 or more	8.8	12.8

The BLS schedules recorded 29 percent supplementary earners and 32 percent nonearners in family units; the Census, 21 percent supplementary earners and 39 percent nonearners. Reconciliation of the differences in definition of the family would exaggerate these differences, since in the BLS Washington study some families of related persons were divided into economic families.

There is little reason to suspect the statistical validity of the sample designs employed by the two agencies. Basically they were the same. Both were ratio subsamples of stratified area

samples designed to obtain population and housing data. Both included persons living in rooming units as well as persons occupying family dwellings. The Census sample, 4,254 families and individuals, was selected from the entire Washington Metropolitan District. The BLS sample, 323 families and individuals, was confined to the Washington area as defined for housing surveys, which is slightly smaller. The difference in area coverage could not be expected to cause significant differences in the results.<sup>1</sup>

Except for the strikingly smaller proportion of individuals covered by the BLS, the sample units drawn by the two agencies are similar.

#### Sample Units, Percentage Distribution by Selected Characteristics

	1944 Urban		1947 Washington, D.C.	
	Census	BLS	Census	BLS
Individuals	21.9	15.9	28.7	15.5
Families	78.1	84.1	71.3	84.5
2 persons	36.7	33.1	35.1	34.8
3 persons	25.6	27.9	26.6	26.0
4 persons	18.7	20.9	20.4	22.3
5 persons	19.0	8.7	10.2	9.9
6+ persons		9.4	7.7	7.0
White			79.2	76.5
Negro			20.8	23.5
D.C.			66.7	69.4
Maryland & Virginia			33.3	30.6

The serious difference in the ratio of individuals to family units cannot be attributed to differences in sample design, but can be explained in large measure by the BLS failure to adjust for nonreporting. Many individuals interviewed by the BLS were classified as ineligible by the economic family definition and the rules for family reconstruction. Among these were students living in the area but belonging to economic families living elsewhere, and persons who were living alone at the time of the interview but who were members of other economic families during the survey year. The proportion of refusals and no contacts was higher among individuals than among families.

<sup>1</sup> 'Significant difference' is used only when the estimated difference exceeds 2 standard errors.

By the substitution procedure explained in the Appendix, the Census adjusted its results for refusals, other nonreporting, and sampling deficiencies, raising the mean and aggregate income about 2 percent. The greatest effect was in the \$6,000 and over group where the number of families and individuals was raised approximately 7 percent. Had BLS applied the same adjusting techniques, the differences in the income distributions would be slightly larger.

## 2 Census and FRB Data

Despite some differences in definitions and methodology, the Census and FRB income distributions were fairly similar. When they differed, the FRB usually had the higher median income and a larger proportion of units in the upper income brackets. A consistent difference, present in all 3 years, was the substantially higher proportion of units with incomes of \$10,000 and over in the FRB distributions (Table 2). It was the chief reason why the FRB income aggregates were higher than the Census.

The distributions of the various subgroups of the population in the 2 sets of samples differ more than the national figures because the data for subgroups are based on fewer cases.

Table 2  
Percentage of Families and Individuals with Incomes of \$10,000 and Over, Census and Federal Reserve Board Distributions, 1945-1947

	1945		1946		1947		FRB adj. <sup>a</sup>
	Census	FRB	Census	FRB	Census	FRB	
United States							
Families & individuals	1.3	2.0	1.9	2.9	2.5	4.1	4.0
Families	1.4	2.3	2.2	3.1	2.7	4.5	4.4
Individuals	0.3	NA	0.4	1.6	1.0	1.3	1.3
Urban <sup>b</sup>							
Families & individuals	1.6	2.5	2.1	3.2	2.8	4.7	4.5
Families	1.8	3.0	2.5	3.6	3.1	5.1	
Individuals	0.4	NA	0.3	0.7	1.1	1.8	
Rural <sup>b</sup>							
Families & individuals	0.7	0.8	1.6	2.4	2.0	3.1	3.1
Families	0.8	0.8	1.7	2.3	2.2	3.4	

<sup>a</sup> FRB data adjusted by reweighting in accordance with Census proportions of urban and rural families and individuals for purposes of methodological comparison.

<sup>b</sup> FRB urban data for 1945 include units living in rural sections of the 11 largest metropolitan areas.

Even though many of the medians and percentage frequencies at given levels are fairly close (Table 3), the differences between some Census and FRB statistics are beyond the limits of probable sampling variability. Differences between the 80 or so pairs of medians and percentages in given income classes in Table 3 were tested for significance. Not all these tests are independent, of course, but they serve to point out the areas where the distributions differ to a statistically significant degree. The number of pairs showing significant differences increased from 3 in 1945 to 7 in 1946 to 14 in 1947. Even after reweighting FRB 1947 data in accordance with Census population weights, 13 significant differences remained. In each case, Census had a lower level of income. The consistency of the direction of these differences is further evidence of a real difference in level between Census and FRB income data. The differences that are beyond the sampling range almost always occur at the very top or bottom of the distribution. Comparisons of Census and FRB data for farms are not possible because the FRB classified income data for the 'farm operator', not for farms as such.

One factor affecting the relative levels of the FRB and Census income distributions was a difference in the proportions of certain subgroups in the samples of the 2 agencies. In all 3 years surveyed, the FRB sample contained a higher proportion of all family units, all urban units, and urban family units, as well as of family units within the rural group. These differences had the effect of raising the level of the FRB income distributions relative to those of the Census. The causes of the differences cannot be precisely determined. Further investigation as well as data forthcoming from the 1950 Census of Population will make possible a more complete examination. However, certain factors affecting the proportion of subgroups can at least be listed:

a) The FRB survey covered only households; the Census covered households in 1945 and both households and quasi-households in 1946 and 1947. In listing dwelling units for interview in the FRB survey some rooming houses may have been omitted because they appeared to be quasi-households. Since rooming houses contain primarily individuals the effect of this possible







Income Class	Families and Individuals			Families			Individuals		
	CB	FRB	FRB adj. <sup>b</sup>	CB	FRB	FRB adj. <sup>b</sup>	CB	FRB	FRB adj. <sup>b</sup>
<i>Medians:</i>									
FRB method, \$		2,920	2,820		3,100	3,080		1,200	1,200
Census method, \$	2,727	2,960	2,860	3,033	3,170	3,140	983	1,330 <sup>d</sup>	1,340 <sup>d</sup>
Difference, FRB-CB, \$		233 <sup>d</sup>	133		137	107		347 <sup>d</sup>	357 <sup>d</sup>
<i>Families &amp; individuals:</i>									
Group %	100.0	100.0	100.0	84.71	88.26 <sup>d</sup>	84.71	15.29	11.74 <sup>d</sup>	15.29
Difference, FRB-CB					3.55 <sup>d</sup>			-3.55 <sup>d</sup>	

*Sampling error of differences between:*<sup>a</sup>

Medians	145			159			314		
Group %				1.8			1.8		

#### Urban

	CB	FRB	FRB adj. <sup>b</sup>	CB	FRB	FRB adj. <sup>b</sup>	CB	FRB	FRB adj. <sup>b</sup>
Under \$1,000	13.0 <sup>d</sup>	7.2 <sup>d</sup>	8.5 <sup>d</sup>	6.3	3.4		44.3	32.2	
1,000-1,999	15.5	15.8	16.4	13.2	14.2		25.7	26.5	
2,000-2,999	21.3	19.8	20.0	22.1	19.5		18.2	22.7	
3,000-3,999	18.9	18.4	18.0	21.5	19.9		6.9	8.8	
4,000-4,999	11.9	13.9	13.3	14.0	15.3		1.9	4.1	
5,000-5,999	7.6	8.9	8.5	9.0	9.9		1.1	1.7	
6,000-9,999	9.0 <sup>d</sup>	11.3 <sup>d</sup>	10.8 <sup>d</sup>	10.8	12.7		0.8	2.2	
10,000 & over	2.8 <sup>d</sup>	4.7 <sup>d</sup>	4.5 <sup>d</sup>	3.1	5.1		1.1	1.8	
Total	100.0	100.0	100.0	100.0	100.0		100.0	100.0	

*Median:*

FRB method, \$		3,280	3,160		3,500			1,620	
Census method, \$	3,009	3,330	3,230	3,350	3,580 <sup>d</sup>		1,208	1,670 <sup>d</sup>	
Difference, FRB-CB, \$		321 <sup>d</sup>	221 <sup>d</sup>		230 <sup>d</sup>			462 <sup>d</sup>	

*Families & individuals:*

Group %	61.69	63.76	61.69	50.85	55.54		10.84	8.22	
Difference, FRB-CB		2.07			4.69			-2.62	

*Sampling error of differences between:*<sup>a</sup>

Medians	202			211			396		
---------	-----	--	--	-----	--	--	-----	--	--

#### Rural

	CB	FRB	FRB adj. <sup>b</sup>	CB	FRB	FRB adj. <sup>b</sup>	CB	FRB	FRB adj. <sup>b</sup>
Under \$1,000	23.0	23.2	24.1	17.3	19.4		66.4	60.0	
1,000-1,999	21.4	22.6	22.8	21.7	22.5		19.4	24.4	
2,000-2,999	20.0	19.1	18.6	21.8	20.0		6.9	8.3	
3,000-3,999	15.6	13.7	13.4	17.1	14.9		3.7	2.3	
4,000-4,999	7.4	7.3	7.1	8.1	7.7		1.9	2.4	
5,000-5,999	5.1	4.3	4.3	5.7	4.6		0.4		
6,000-9,999	5.5	6.7	6.6	6.1	7.5		0.6	2.6	
10,000 & over	2.0	3.1	3.1	2.2	3.4		0.7		
Total	100.0	100.0	100.0	100.0	100.0		100.0	100.0	

*Medians:*

FRB method, \$		2,150	2,100		2,350			810	
Census method, \$	2,252	2,220	2,170	2,450	2,410		645	830	
Difference, FRB-CB		-32	-82		-40			185	

*Families & individuals:*

Group %	38.31	36.24	38.31	33.86	32.72		4.45	3.52	
Difference, FRB-CB		-2.07			-1.14			-0.93	

<sup>a</sup> FRB urban data for 1945 include units living in rural sections of the 11 largest metropolitan areas.

<sup>b</sup> FRB data adjusted by reweighting in accordance with Census proportions of urban and rural families and individuals for purposes of methodological comparison.

<sup>c</sup> Five percent level of probability.

<sup>d</sup> Difference larger than sampling error of difference.

omission, most likely in urban areas, would be to understate the number of secondary individuals.

b) Different definitions of the dwelling unit made the FRB exclude and the Census include some rooming houses and semi-apartment hotels among households. Since it is individuals who usually live in such quarters and they in turn are essentially a characteristic of cities, this probably tended to increase the proportion of both primary and secondary individuals in the Census sample relative to the FRB sample, and also the proportion of urban units.

c) The Census probably treated most unmarried children absent at college who were mainly self-supporting as family members, while the FRB treated those living in households as individuals and excluded completely those living in quasi-households, e.g., dormitories and large rooming houses. This would tend to increase the proportion of individuals in the FRB sample.

d) The Census, but not the FRB, adjusted the number of individuals in the 1946 and 1947 income samples to certain predetermined population totals with a specific age-sex distribution (see App. B). Certain types of persons, such as young single males, are especially difficult to find at home, and may therefore be omitted by the enumerator. While the Census adjustment to predetermined population totals has merit as a corrective for sampling variation for the entire population it may yield a slightly too high proportion of individuals in private households. Since the FRB did not use outside controlling estimates of population, its samples may contain too few of the primary and secondary individuals who are difficult to find at home.

e) There is some evidence that the FRB sample points for 1945 and 1946, which were randomly selected from stratifications of sampling points, were slightly more urban than the strata from which they were taken, tending to yield too high a proportion of FRB urban units. For 1947 FRB interviews were weighted to correct for this factor.

Why were differences between the Census and FRB income distributions much less marked for 1945 than for 1946 and 1947

when the FRB showed consistently higher levels of income? Our hypotheses can be only tentative at best and, unfortunately, are sometimes based on little more than personal judgment. In the absence of controlled experimental evidence it is extremely difficult to appraise cause and effect in a complex situation.

The reason for the greater disparity in later years was certainly not an increase in the proportion of certain population groups, such as urban units, especially urban families in the FRB sample with above-average incomes. In fact, somewhat of a reverse tendency may be noted. The proportion of urban families in the FRB sample decreased slightly relative to the Census proportion from 1945 to 1947. Further, the proportion of family and individual rural groups, which are below-average in income, increased in the FRB sample more than in the Census sample.

This point may be demonstrated in another way. Reweighting the 1947 FRB income data for 4 groups (urban families, urban individuals, rural families, and rural individuals) in accordance with Census survey proportions tended to lower the median value of combinations of these 4 groups. However, only one of the significant differences between FRB and Census medians was reduced sufficiently to bring it within the range of sampling error. Even in this one case, the FRB median for all families and individuals was \$133 higher after reweighting than the Census median, just within the \$145 range of sampling error. The FRB medians for all individuals and for urban families and individuals were still significantly higher after reweighting than the respective Census medians.

The changed relation between the Census and FRB income distributions appears to be due partly to differences in definitions and timing of the surveys. The chief factor, however, is thought to be methodological. Both agencies made at least one basic change in their methodology during the 3 years (see App. B). The FRB asked for fewer details on income in 1945 than in 1946 or 1947, while the Census asked for fewer details in 1947 than in 1945 or 1946. Apart from special factors, then, the data for these 3 years may be examined with a view to establishing the relative level of income yielded by the respective general

methodologies, and the use of a few versus many income questions.

For 1945, the year of small detail for FRB and great detail for Census, the 2 distributions were very similar. However, differences between 3 of the pairs of figures in Table 3 were large enough to be statistically significant. In all 3 the FRB had the significantly higher figure, that is, a larger proportion at the top of the income distribution or a smaller proportion at the bottom.

For 1946, the year of great detail for both agencies, the level of the FRB distribution was generally somewhat higher except for certain rural distributions, though most of the differences were too small to be statistically significant. However, the significant differences in medians as well as proportions increased to 7 in 1946 and again the FRB had the significantly higher income figure.

For 1947, the year of small detail for Census and of great detail for FRB, the level of the FRB distributions was generally higher. The significant differences increased to 14. Even after FRB data were reweighted in accordance with Census population weights, 13 significant differences remained. Again, FRB had the higher level of income in each case of significant difference.

Summarizing: in 1945, when the FRB had less income detail, the levels were approximately the same, although there were some evidences of a higher FRB distribution. In 1946, when both agencies had more detail, the nonsignificant differences were somewhat larger and the number of significant differences doubled. In 1947, when the Census had less detail, the level of its distribution was substantially below the FRB and the number of significant differences doubled once again.

The above facts would seem to indicate that the number of questions may have a substantial effect upon the level of the distribution, especially when the survey is devoted solely to income or when financial information other than income is not sought. However, surveys that encompass other financial items may not be as sensitive to the expansion or contraction of specifically income questions, because their other financial questions

may sometimes serve as memory prods. In some cases, the other financial data may serve as a check also on inconsistencies in the income data.

The conclusion that the FRB distributions tended to be somewhat higher at some points, even for 1945, is further supported by evidence of a special factor that tended to raise the median Census incomes in 1945 an estimated \$50–100 relative to those of the FRB: the difference in the timing of the surveys in a period when the armed services were demobilizing rapidly. The Census survey took place during the week including April 15, 1946, the FRB survey almost wholly during January and February. Between February 1, the average date for the FRB survey, and April 15, 2 million persons with an average 1945 income of about \$1,000 were released from the armed services. Well over half probably became members of multi-person units. Most of these added the balance of their incomes after allotments, including military pay, to the incomes of the other members. The difference in the timing of the surveys, therefore, should properly have made the Census median income for 1945 higher than, rather than similar to, the FRB figure.

What factors other than income detail may explain why the FRB income figures were higher than those of the Census Bureau? Possibly the chief factor was that the information was supplied by the actual earner, not the housewife, in about three-fourths of the cases in the FRB surveys and by only an estimated fourth to half in the Census surveys. Housewives may not know about extra earnings in the form of commissions, bonuses, overtime, etc.; they may be familiar only with take-home pay and not know about income before deductions. Also, in some cases, they may not have been told about a husband's or a child's increases in pay. Moreover, they have less contact than the earner with income tax forms or annual earnings statements (Form W-2) which are excellent memory prodders.

Another factor was the long conversations with FRB interviewers covering many financial transactions apart from income which brought to mind forgotten income items and were a better check on consistency than the population, housing, and labor

force information collected by the Census. Introductory questions and others spaced throughout the interview were designed to give the respondent opportunity to express himself and not just answer yes or no or supply a missing figure. It is not yet established whether this more informal type of interview yields definitely better results. However, some increment of information might well be expected from an interview in which a special effort is made to foster cooperation, interest, and responses, and other financial items are discussed. This increment might be present at all levels of the income distribution, and might, perhaps, be largest at the higher levels, especially when units are reluctant to give information or from which information is most difficult to get because the income is from several sources.

An additional factor was the more efficient FRB system of stratification which permitted a more precise treatment of units for whom information on income was not collected. It is widely accepted that the lowest interview completion rate in income surveys is that for the highest income group. Isolation of this group fairly precisely in a relatively homogeneous stratum makes it possible to increase the weights of the reporting members to compensate for those in the stratum from whom income data could not be obtained. The comparable Census technique for handling such cases, which assumed that the nonreporting cases had the same incomes as the part of the sample that reported, ignored the much higher rate of nonreporting by upper income units. The effect was to understate the proportions of units at the highest levels.

Some differences are due to differences in the definition of the family unit but their net effect is believed to have been small. Other factors were the FRB acceptance of gross rental income figures from heads of households with fewer than 4 roomers and boarders, and from persons who were renting part of their own home or apartment, while the Census calculated net incomes in both cases. Also, the FRB included the income of a family member who died during the income year or just before the interview, in the few cases when the information was volunteered.

### 3 1946 Census-BAE Rural Farm Data

The primary reason for excluding 1946 rural farm data, covered by the BAE, from the published Census survey tables was the unusually large proportion of farm families and individuals in the BAE sample reporting a loss from farm operations. In the 1944 and 1945 Census samples 7.6 and 9.2 percent respectively of families and individuals with farm income or loss reported a loss from farm self-employment; in 1946, according to the BAE data, approximately 33 percent.

Census and BAE results can be compared more directly, though in a limited way, because both agencies covered families and individuals on farms of less than 10 acres in 1946. Slightly over 72.5 percent of this group reported a loss from farm operations in the BAE sample, 2.7 percent in the Census sample; the corresponding figures in the \$1-499 class were 17.6 and 68.6 percent, respectively; in the \$500-999 class, 4.1 and 12.2 percent; and in the \$1,000 and over class, 5.7 and 16.5 percent.

The BAE figures for this group, about 400,000 as estimated from tabulations of the sample, represented over 100,000 more families and individuals than the Census figures. The 1945 Census of Agriculture reported almost 600,000 farms under 10 acres. The difference can be attributed at least in part to sampling variation and the Census exclusion of urban farms. However, it cannot have affected the percentage distribution sufficiently to account for the difference between the BAE and the Census figures. Some of the difference may be attributed, of course, to sampling variation, but it seems certain that other factors also were present; e.g., differences in the classification of farms. The differences are not due to differences in definitions, for the Census and BAE definitions of a farm and also of income were meant to be identical. The main explanation may be differences in the techniques used to collect income data together with the difficulty of identifying small farms.

The Census sample figure, 3 percent in 1946 for families and individuals operating farms under 10 acres who reported a loss, seems to be much too low compared with those for farm families obtained in the 1944 and 1945 Census samples for all sizes of farms combined, 7.6 and 9.2 percent, and compared with that

for farms of 10 or more acres in the 1946 BAE sample, 27 percent. On the other hand, the BAE loss figure for farms under 10 acres, 73 percent, and the BAE loss figure for all farms, 33 percent, seem at first to be unreasonably high. However, a large number of losses are to be expected when the definition of farms, which embraces many small scale part-time operations, is taken into account. Agricultural tracts of 3 acres or more or yielding products, those consumed at home as well as sold, valued at \$250 or more were included. In the 1945 Census of Agriculture about a third of all farms in the United States had gross sales of less than \$500 in 1944, and 10 percent reported no sales. This provides some substantiation for a high loss rate.

For farms under 10 acres data by size of gross sales are not available from the Census of Agriculture, but since over half reported less than \$500 worth of products consumed at home or sold, the proportion with gross sales of less than that amount was doubtless appreciably higher. As the value of home-produced home-consumed farm products was not in the net cash income concept used in any survey, a large proportion of net cash losses from farm operations is to be expected.

Net income also was understated because production expenses included certain items such as taxes and similar expenses on the farm residence which are not true farm production expenses since the imputed income from farm dwellings is not included in gross receipts. This undoubtedly understated net income, at least at the lower levels, more than the exclusion of depreciation charges overstated it. The understatement could be expected to produce a large number of farm losses among families and individuals on small farms. For the BAE sample of farms of 10 acres or more, however, the 1946 losses are still 27 percent, and for farms of 500 acres or more, about 23 percent. The high loss rate for the large farms may in part be due to the exclusion of inventory changes from the definition of income.

Differences in the type of farm schedule probably explain the major part of the differences in the farm loss rate. As Mr. Koffsky and Miss Lear show in Part V, the 1946 reports to the BAE on gross farm receipts, like similar reports to the Bureau of the



Census for the 1945 Census of Agriculture, were gross understatements, whereas the reports to the BAE on farm expenses were reasonably close to maximum estimates. One can only hazard a guess why the Census global approach in 1946 yielded fewer losses than the BAE detailed approach. One possible explanation is that whereas farmers understated only gross receipts when reporting in detail they omitted certain expenses also when reporting to the Census. Although the Census and BAE definitions of receipts and expenses were identical, we do not know whether the same items were reported, for the Census schedule did not list them. It seems likely, for example, that some families and individuals in the Census farm sample did not think of deducting all taxes, including taxes on the residence, from the gross receipts of what in reality were large suburban gardens rather than professional farm enterprises but qualified for classification as farms. Also many expenses associated with the farm garden were doubtless not reported except at the enumerator's probing.

Another factor in explaining the discrepancy may be the high rate of nonreporting and incomplete reporting of total money income in the BAE survey, more than 30 percent of the identified farms, i.e., all those on the sample lists of addresses to be visited. When inflated, the number of identified farms fell short of the 1945 Census of Agriculture count. The true rate of nonreporting and incomplete reporting of total income, therefore, somewhat exceeded 30 percent. In their study of income from farm operations Mr. Koffsky and Miss Lear were able to use many schedules that had been discarded by the Census because nonfarm items were missing. Furthermore, for the gross income portion of their estimates they were able to use a larger sample because total income data were requested of a subsample. The greater representativeness, in general, and the lower nonreporting rate, in particular, of their schedules were reflected in changes in the distribution of net farm income, but these changes did not greatly affect the large discrepancy now under consideration.

In discussing the BAE reporting rate above we referred to the

undercount of farms. The totals of farm families and individuals reported in the 1944 and 1945 Census surveys, approximately 4,600,000 and 4,500,000 respectively, also were too low compared with the approximately 5,900,000 farms enumerated in the 1945 Census of Agriculture. The 1946 Census-BAE sample estimate of the number of families and individuals with farm income or loss, 6,045,000, appeared to differ from the 1945 Census of Agriculture number of farms by no more than should be expected. Although the number of persons receiving income (or loss) from farming in 1946 agreed approximately with the number of farms reported in the Census of Agriculture, no specific adjustment was made to compensate for the failure to identify many small farms, particularly those under 10 acres. Mrs. Goldsmith estimates that if the number of families and individuals with farm income had been reported as accurately in 1944 and 1945 as in 1946 and the greater accuracy in reporting the number had not affected the percentage size distribution (though the chances are that it would have done so), the proportion of aggregate net farm income in the inflated survey would have been about the same in 1945 as in 1946 and higher in 1944 than in 1946. In other words, the higher mean net farm income in 1944 as computed from Census data was more nearly accurate than the 1945 or 1946 income (Table 4), but not necessarily the distribution.

Why did the Census-BAE 1946 sample yield a higher mean farm income than the Census 1945 sample despite the much larger proportion of losses and what appears to be a lower median net income from farm self-employment?<sup>2</sup> The increase in losses seems to have been at the expense of the next higher net farm income level, \$1-999. The percentage earning less than \$1,000 did not increase significantly from 1944 and 1945, when the figures were based on Census data, to 1946, when the figures were based largely on BAE data, but the percentage earn-

<sup>2</sup> Although the median net income from farm self-employment in Table 4 is \$589 for 1945 when based entirely on Census data and \$516 for 1946 when based mainly on BAE data, the difference is small enough to be due to sampling variation. The large increase in estimated aggregate net farm income, however, should lead us to expect a definite increase in the median.

ing \$3,000 or more increased, as would be expected, with the actual increase in farm income.

Table 4

Rural Farm Families and Individuals, Percentage Distribution by Class of Net Income from Farm Self-employment, 1946, 1945, 1944

<i>Income Class</i>	<i>1946<sup>a</sup></i>	<i>1945</i>	<i>1944</i>
Loss	26.1	8.9	7.3
\$1-499	23.5	37.9	37.0
500-999	14.9	16.9	19.3
Total under \$1,000	64.5	63.7	63.6
\$1,000-1,999	14.6	18.9	18.7
2,000-2,999	8.4	10.4	9.0
3,000-3,999	4.6	2.5	3.7
4,000 & over	7.9	4.5	5.0
Total reporting net farm income or loss	100.0	100.0	100.0
Net farm inc. for fam. and indiv. reporting net farm inc. or loss (\$)			
Median	516	589	644
Mean <sup>b</sup>	1,188	1,032	1,398

<sup>a</sup> Includes both Census and BAE cards. Percentage in loss level would have been higher if only BAE returns had been used. The substitution of Census returns for families and individuals on farms under 10 acres for the comparable BAE group lowered the loss rate, for only about 3 percent of the Census families and individuals in this group reported a loss compared with 73 percent of the BAE families and individuals.

<sup>b</sup> Includes estimates by Mrs. Goldsmith for the \$10,000 and over income class; see Part VI, Table 5.

Probably some of the factors that may have caused a large number of losses on the BAE returns had relatively more influence on the net income of farms with small gross incomes than of farms with larger gross incomes, thus to some extent making for more inequality in the BAE than in the Census distribution. For example, the deduction from gross farm income of taxes on the farm dwelling might have caused a large number of farms with small gross incomes to have losses, but might have had little appreciable effect on the net incomes of farms with large gross incomes. Furthermore, there is no reason to suppose that the detailed listing of receipts would tend to produce much better reporting for farms that have small gross receipts and sell only 1 or 2 items, but it may very well account for items that would otherwise be missed for farms carrying on large and diverse operations.

We have compared the net farm income estimates based on BAE returns with those based on Census returns, but we have not yet compared the results of the 2 sources of data with respect to total money income. According to estimates based on the BAE sample, 8 percent of the farm-operator families and individuals on farms under 10 acres suffered a loss in total money income; 2.5 percent of Census families; 42 and 38 percent respectively had a total money income of \$2,000 or more. Thus for this group of families and individuals operating small farms, whose total money income is apparently from predominantly nonfarm sources, it may be inferred that the estimates of income from nonfarm sources based on BAE data were as high as or higher than the estimates based on Census data. Although the proportion of all farm families and individuals with money income from sources other than farm self-employment is large—more than half according to estimates based on 1944 Census data—the reporting of the nonfarm income component would not have as much effect on the total money income of all farm-operator families and individuals as on the total money income of those residing on farms under 10 acres.

The estimated median total money income of rural farm families and individuals in 1944 (\$1,157), 1945 (\$1,291), and 1946 (\$1,275), like net income from farm self-employment, remained constant from 1945 to 1946, a period during which an increase is to be expected. On the basis of the evidence we have considered it seems reasonable to believe that this stability of the figures is due to modifications in survey procedures (see App. Table 2 for rural farm and rural nonfarm size distributions 1944-47).

### C FIELD SURVEYS AS A TECHNIQUE FOR CONSTRUCTING INCOME SIZE DISTRIBUTIONS

When compared with independent estimates of aggregate income and income size distributions, field surveys seem to have consistently understated consumer income. The underestimates have varied from one year to another, from one type of income

to another, and from one type of survey to another. There is no assurance that the effect of these variations in survey efficiency on the unadjusted size distributions is not to exaggerate, obscure, or do both to real changes in the nature of the size distribution. However, the survey figures, an essential component of the various sets of data used in constructing adjusted income size distributions, are the one source for relatively quick preliminary estimates. The collation of field survey data with those from such other sources as income tax returns, payroll tax data, and estimates of aggregate income of various types is still in the experimental stage; moreover, it requires considerable time. Before we can know to what extent and how variations in efficiency qualify inferences drawn solely from survey data, further research into both field surveys and the adjustment process is necessary. It is not clear whether the qualifications that must be attached to the chief uses of field survey data are minor or so substantial as seriously to affect their value in unadjusted form.

The knowledge gained through diverse approaches to field collection of income data doubtless affords a basis for improving techniques and procedures of income surveys. It is believed, however, that all data collected through field surveys by the agencies considered in this report would be much more valuable and better procedures more readily developed if the agencies would supplement their reports by publishing data based upon common concepts and definitions so that all the results for a given year could be compared. For this reason it is desirable that field surveys of income be coordinated, at least with respect to standard definitions of income and the income-receiving unit and techniques such as family reconstruction. As estimates of aggregate income and its size distribution are improved by continued effort to bring together all information on consumer income, probable as well as known deficiencies should be stated so that users can readily assess the data. Basically, the agencies conducting field surveys of income used the same sampling methods—a ratio subsampling of area samples based on accepted probability sampling theory. Variations in sampling techniques, e.g., Census selection of clusters, probably have little

effect on the sampling variability. The variation inherent in sampling explains some of the differences in results. However, it does not account for all the differences, some of which are serious enough to warrant continued study of survey techniques and a determined effort to improve them.

### 1 *Universe*

The universe covered by each survey was dictated by the major objectives and the funds and facilities available. The BLS studies were confined to urban areas; the BAE to the farm population; FRB studies to households (to make the most efficient use of funds for the collection of consumer financial data); Census studies covered populations ranging from persons in households only to the entire civilian noninstitutional population of the nation. Data obtained in connection with each study are a valuable source of information about consumer income, but they must be pieced together and studied before they tell a complete national story.

### 2 *Income-receiving Unit*

The recipient unit, like the universe covered, is designed to satisfy the major objectives of field surveys. The BLS economic family seems appropriate for studies of the distribution of expenditures for current consumption by income class; the FRB spending unit is suited to the study of consumers' liquid assets and their intentions to purchase consumer durable goods; the Census family is appropriate for population surveys.

However, the use of the income size distribution in studies on income inequality, income level, poverty, income parity, and the consumption functions requires that definitions of the recipient unit fit the purposes and general needs for income data in relation to the economic phenomenon under study. A definition of the income-receiving unit that depends upon the use of income and the degree of dependence of individuals in the family upon the head might be appropriate in most instances. Some such definition should be developed and adopted as a standard unit; and in future field surveys the information on schedules should be sufficient for income data, per se, to be summarized on the

standard unit basis. Until a standard unit is adopted, the Census definition might well serve as a standard by which changes in the size distribution of income and the efficiency of future surveys can be judged, especially if the decennial Censuses are to be exploited to their maximum potential.

### *3. Family Reconstruction*

We believe that taking the composition of the family as of the time of interview may, during periods of rapid change, qualify the data to such a degree as to invalidate their use for some important purposes; that the additional survey cost of reconstruction, due mainly to the greater complexity of editing and coding about a fifth of the schedules, should be absorbed in some manner to make reconstruction possible.

Schedules from past surveys do not contain the information that would make possible a realignment of the family membership and an adjustment of family incomes for comparative purposes. It is of interest, however, to consider the possible differences in results the two procedures might produce.

a) In some cases failure to reconstruct the family tends to increase the number of families at lower income levels.

- 1 Marriage between 2 members of separate families who form a new household
- 2 Divorces
- 3 Undoubling of related families
- 4 Family members entering the armed forces
- 5 Family members leaving home to become self-supporting and live apart
- 6 Family members going abroad or entering institutions
- 7 Deaths

In each case the total income of the family as originally constituted is reduced by the income of each member who leaves the family. Except when a family member joins the armed forces, goes abroad, enters an institution, or dies, a new family or single person unit is created, and the same aggregate income is distributed among more units.

b) In other cases failure to reconstruct the family tends to increase the number of families at higher income levels.

- 1 Marriage between individuals who are not members of families
- 2 Doubling up of related families
- 3 Veterans returning from the armed forces
- 4 Self-supporting individuals who have lived alone and join a family through marriage or otherwise
- 5 Family member returning from abroad or from an institution

In each case, except for the returning veteran or other returning family member, families or individuals are merged and the same aggregate income is distributed among fewer units. When a veteran or other family member returns, the family income is increased by the amount of his earnings.

c) In a few cases failure to reconstruct the family may have little effect upon the income size distribution.

- 1 Marriage between an individual who has lived alone and a family member who establish a new household
- 2 Any realignment of families or individuals that produces the same number of units as existed before the rearrangement

In these cases, since the same aggregate income is distributed among the same number of units the average or median does not change.

The difference between income size distributions of BLS reconstructed and nonreconstructed families for any one year due to a combination of the divergent effects of these possible family changes depends upon the economic situation. During periods of prosperity and high employment families may be expected to undouble as fast as housing becomes available; children to leave home to become self-supporting; newly married couples to form new households; and the divorce rate to rise. Failure to reconstruct the family in surveys conducted in prosperity would tend, therefore, to increase the number of families



in the lower income classes and lower the median income. The converse may be expected during periods of depression and unemployment when families are doubling up to help meet living costs; sons and daughters, with their means of self-support lost through unemployment, return to the family; and the son- or daughter-in-law comes to live with the family. Failure to reconstruct the family seems to bias the size distribution downward when incomes are increasing and upward when incomes are declining, except to the extent that undoubling is prevented by a housing shortage.

In times of mass induction into the armed forces, the income distribution based on nonreconstructed families will be biased downward; when the armed forces are being demobilized, the bias will be upward. In a stable economy these divergent effects of changes in family composition might well maintain some balance. This cannot be determined without further study.

These effects of failure to reconstruct the family are in the same direction whether its composition changes during or after the survey year. However, since the BLS includes in total family income only the income received by individuals while members (part-year), the effect is more serious the nearer the date of the change in composition is to the year-end. If the composition changes after the year-end the seriousness of the effect grows the longer the interval between the year-end and the interview, and the chance of change increases.

In area studies, such as those conducted by the BLS and Census in Washington, D.C. for 1947, and by the BLS in selected cities for 1945 and subsequent years, families that leave or move into the area during or after the survey year but before the interview pose still another problem of reconstruction. Families that leave the area cannot of course be counted. Those that come into the area will be treated differently depending upon the thoroughness of the reconstruction of the family.

Failure to recognize changes in the area population during and after the survey year may introduce biases in survey results, especially when the population is growing, as happened in some urban areas during the war. If the total annual income of

families that migrated to the area is included, the data will not be representative of the area population in the survey year, and the bias in the distribution of incomes will depend upon the income level of the in-migrating population. When the population of an area is stable, inclusion of the income of in-migrants offsets in some degree the loss of incomes received by out-migrants. In the wage earner surveys of 1934–36 the BLS excluded families that had resided in an area fewer than 9 months, but in its more recent area studies this rule was not applied.

The BLS procedure is, of course, only one of several techniques that might be devised to solve this problem. Each technique will yield different results. For example, if the reconstruction is based upon fractional parts of the year during which the composition did not change, and the income data are summarized by fractions of income-receiving units at annual rates of income, the distribution will be quite different from the distribution obtained by the BLS technique. To illustrate, let us consider a marriage on October 1 between a self-supporting woman and a man who was a member of a 3-person economic family before marriage. Assume that the wife, husband, and the head of the 3-person family each earns \$4,000 a year. The BLS reconstructed-family technique will yield 2 income-receiving units: 1 at \$7,000 (total income of the head and 75 percent of the husband's income) and 1 at \$5,000 (total income of the wife and 25 percent of the husband's income). The fractional unit technique will yield 75 percent of a unit at \$4,000 (75 percent of a unit at the annual income of the wife) and 1 at \$8,000 (75 percent of a unit at the annual income of the head and husband, and 25 percent of a unit at the annual income of the husband and wife). If the husband was promoted to a \$6,000 per year job at the time of his marriage, the number of units would then be 75 percent of a unit at \$4,000 (75 percent of a unit at the annual income of the wife), 75 percent of a unit at \$8,000 (head and husband before the latter's marriage), and 25 percent of a unit at \$10,000 (husband at \$6,000 and wife at \$4,000 per annum). We believe that such a method would yield a more accurate picture of the distribution of income-receiving units throughout the

survey year. However, it might prove difficult or impossible to apply to expenditure and savings data. The value of field surveys as a source of income data may be greatly enhanced by further exploration along these lines.

#### *4 Date of Interview*

The increase in the error of recall on the part of the respondent and the reduction in the accuracy of the reported data is probably most serious when data are collected for reconstructed families and the part-year members are no longer present. On the other hand, there is a decided advantage in delaying the collection of income data until wage and salary earners have received Form W-2, and can refer to their income tax returns when answering income questions. W-2 forms are ordinarily in the hands of a substantial proportion of all earners by January 15, and are received by nearly all earners by the end of January. Most Form 1040 returns, however, are not filed until March. For this reason, income from nonfarm self-employment might be reported more accurately in April.

When no attempt is made to reconstruct the family, the lapse of time between the year-end and interview assumes importance with respect to the biases due to failure to reconstruct. The longer the interval the greater the bias.

The BAE surveys and the Census Washington, D.C. survey were begun soonest after the end of the survey year. Most of the FRB data on liquid assets and consumer finances have been collected in January and February. Census data for 1944 were gathered in May, and for the following three years in April. BLS data for its 1941 study of consumer expenditures were collected in April-June; the 1944 information was collected in January-April; BLS studies have since run in the field from February to as late as April, though the majority of the survey schedules were completed in February and March.

#### *5 Income Reporting*

Estimates of the income distribution based on field surveys differ from the true figures for two basic reasons: sampling variability,

discussed below, and biases due to error of response and to non-reporting.

Errors of response occur mainly because the schedule entries for income are usually based on memory rather than on records, and in some instances on the memory or knowledge of some person, such as the wife of the household head, who may not have complete knowledge of her family's income. The memory factor may yield responses that have two possible types of variation: bias, either upward or downward, and random errors. In a fairly large sample, random errors tend to cancel out without producing a bias. The memory factor probably produces underestimates because respondents forget irregular items received by the principal earner. The bias due to the memory factor no doubt increases as a longer period elapses between the year-end and the interview. Income items are doubtless frequently not recorded because the respondent was not the recipient and erroneously supposed that they were not received. The several agencies vary considerably in the degree to which the actual income recipients are the source of the information. Roughly, income recipients were the source in more than a fourth but fewer than half of the Census and Census-BAE cases, in about half of the BLS cases, and about three-fourths of the FRB cases. Other errors of reporting are due to wilful misrepresentation or to misunderstanding on the part of the respondent concerning the scope of the income concept. Wage and salary earners, for example, sometimes cannot report gross earnings and consider take-home pay as their total income. This may occur even more frequently when the information is supplied by persons other than the actual earner, e.g., a housewife, who may know only the take-home pay of her husband or son. The more the memory of the respondent is stimulated through detailed and probe questions, and through direct interview of each recipient the better the data.

### 6 *Nonreporting*

Nonreporting, i.e., the clearly recognizable failure to obtain any or all income information due to refusal, absence, sickness in the

unit, language difficulties, or inability of the informant to provide information, may occur at any income level. If it occurred in the same proportion at all points of the income distribution, the distribution would not be distorted solely on its account. However, in view of the considerable evidence that nonreporting due to refusal to participate in a survey is more prevalent at upper income levels, the estimated proportion of cases at these levels is probably smaller than it should be.

A Census tabulation of the characteristics of the families and individuals in households reporting on family composition and other information but not on one or more income items for 1945 indicates in general that the categories with the highest incomes were those with the highest nonreporting rates. Census nonreporting was at its highest rate in 1945, approximately 20 percent. White families failed to report more frequently than nonwhite; owner families more frequently than tenant; among tenant families those paying high rents more frequently than those paying low rents; and among civilian earners those following occupations in which median civilian money earnings were high tended to have a higher nonreporting rate than those following occupations in which civilian money earnings were low. There were some exceptions; for example, the median total money incomes of families and individuals in places of a million and more inhabitants was higher than that of families and individuals in smaller places, but the nonreporting rate in the former was lower.

What is the net effect on the distribution of income? To answer this question, the Census median income for families and individuals in 1945 was recomputed:

- a) The nonreporting cases were distributed by income level within each monthly rental level by the proportions for the reporting cases within each monthly rental level. The number in the assumed distributions by income level of the nonreporting cases was then added to the number for the reporting cases, and a new median income was computed for the total of all primary tenant families and individuals reporting monthly rent paid.
- b) The nonreporting cases were similarly distributed by income

level within each size of place of residence classification to obtain a new distribution and median for all urban families and individuals.

c) Nonreporting white and nonwhite families and individuals were distributed by income level within each color group on the basis of the percentage distribution of those reporting income. The white and nonwhite columns were then combined, and a new distribution and median computed for all families and individuals.

d) Male civilian earners not reporting earnings were distributed by civilian money earnings level within each major occupation group on the basis of the percentages for those reporting civilian money earnings. The occupation groups were then combined, and a new distribution and median computed for all civilian earners reporting an occupation.

The maximum increase in a median was \$17. In the size of place computation, the median decreased \$4. The effect of the adjustment on the income size distribution was greatest at the \$6,000 and over income level, which was raised approximately 11 percent. Recomputations of 1947 Census medians, made in a similar manner by using nonreporting cases classified by color and urban-rural area, yielded approximately the same differences as for 1945.

The sample income distribution resulting from the 1941 BLS survey was adjusted to take account of nonreporting and substitutions (7 percent) by applying to the distributions of all originally assigned and completed schedules an adjustment factor for the estimated rate of reply at each income level. The adjustment factors were based on nonreporting rates in large city blocks and small cities with different median rent and rental value ranges. Here also the nonreporting rate tended to be higher in blocks with higher rent levels and with larger proportions of families at upper income levels, ranging from about 1 percent at the under \$1,000 level to 35 percent at the \$10,000 and over level. This adjustment increased the median of the distribution of urban families and single consumers about \$35 and changed the shape of the income size distribution substantially

at the lower and upper income levels. In the 1944 BLS survey the substitution rate was about 12 percent (see BLS *Bulletin* 822). However, the unadjusted income distribution was published, and the BLS has not since attempted to adjust for non-reporting.

Rough calculations, which should be considered as minimum estimates of the effect of the FRB stratified treatment of the nonreporting cases, indicate that for the 1947 data the median income was raised about \$50. The effect on the distribution at each level was not large, but the increases were biggest at the highest income levels. Units with incomes of \$10,000 and over increased approximately 7 percent. The FRB system of stratification appears to be most effective in improving the survey data at the highest level.

Clearly some adjustment of survey data for the downward bias due to nonreporting is essential. It seems reasonable to assume that the various techniques used by the several agencies are effective, in varying degree, though none eliminates the entire error. The adjustment procedures can probably be improved by correlating income with additional characteristics, e.g., family size, number of earners, occupation, race, age of head, and ownership of an automobile, which can usually be ascertained for most income-receiving units.

The above computations do not, of course, measure precisely the degree to which nonreporting affects the accuracy of the income estimates. Their assumption, that the nonreporting cases within a specific occupation, rent, or census tract group have the same income distribution as those in the same group who report their incomes, doubtless causes an understatement of the error due to nonreporting.

Mrs. Goldsmith shows in Table 2 of Part VI that aggregate income estimates based on the Census income distributions are only three-fourths to four-fifths of the amounts that should be covered. The missing amount is too large to be accounted for entirely by biases introduced by nonreporting even though the estimates of the latter in the above paragraphs are admittedly understatements. We believe that the deficiency shown by Mrs.

Goldsmith is in considerable degree explained by the respondents' own underestimates. This suspicion is strengthened by our analysis of the differences in United States urban income distributions for 1944 and Washington, D.C. area distributions for 1947, and a comparison of the procedures used in the Census and BLS surveys.

Students have long been aware in a general way that income size distributions based solely on field surveys understate median and mean income and the proportion of units at upper income levels. They have often consoled themselves with the thought that the biases were probably reasonably constant from year to year. Mrs. Goldsmith's Table 3 deprives them of this solace. For several large income components the reporting deficiency varied considerably from year to year. One outstanding instance was the proportion of nonfarm self-employment income reported, which in Census surveys fell from 78 percent in 1944 to 51 percent in 1945, and in 1946 increased to 56 percent of the aggregate derived from the NID personal income series. Another example is the proportion of aggregate interest and dividends reported, which fell in the Census surveys from 32 percent in 1944 to 18 percent in 1945, and was 21 percent in 1946. The proportion of farm income covered also varied considerably. Comparable BLS and FRB figures are not available at present. Although sampling variations may account for some differences, it seems unlikely that they constitute the sole explanation.

Inasmuch as the proportions of total income contributed by some of its major components vary from one income level to another, the instability of the relative accuracy with which the various components are reported may affect the bias in the income distribution. The differing degrees of accuracy with which the various components are estimated during a given year, and the variation in the accuracy with which a given component is estimated from year to year, directly affect the validity of at least two types of analysis. Obviously an analysis of the number and percentage of families and individuals receiving specific kinds of income, such as wages and salaries and profits from self-employment, on the basis of unadjusted field survey data is unlikely to lead to reasonably sound conclusions. Moreover, com-



parisons of income from nonfarm or farm self-employment with income from other occupations are not to be recommended unless the probable accuracy of the figures is rather definitely and quantitatively stated.

Little of the year to year change in the proportion of aggregate income reported can be attributed to changes in the nonreporting rate. As we have seen, the methods used to adjust for nonreporting leave field survey data seriously biased from the viewpoint of average values and perhaps even of year to year comparisons. There is apparently little relation between nonreporting rates and the accuracy of the estimates. The Census nonreporting rates for 1944 and 1945 differed only slightly, 17 and 20 percent respectively, but that for 1945 was double that for 1946. The slight decrease in reporting from 1944 to 1945 was accompanied by a small decline in the percentage of income reported (from 79 to 74) while the great improvement in 1946 reporting over that for 1945 was accompanied by an increase of the same amount (5 percentage points) in the percentage of income reported. In other words, the percentage of income reported for 1944, 79, was the same as for 1946 when the nonreporting rate was substantially lower.

Although a very high nonreporting rate is obviously undesirable because of the possibility of large biases, the evidence indicates that the reduction of the rate leaves the main part of the problem of understated income untouched.

#### *7 Census 1946 and 1947 Inflation Techniques*

The final revisions of the 1946 family and individual weights (App. B) affected the medians only slightly. The maximum difference in the family and individual tables was a \$26 decrease in the median total money income of urban families. The 1947 adjustments (App. B) increased the weights for primary families and individuals (or households) an average of 1.7 percent, for all families and individuals 2.3 percent, and for persons 14 and older 3.2 percent. The effect of some preliminary adjustments, of which the substitution of complete for incomplete returns was one, is not included.

The effect of the adjustments of the 1947 income sample

weights cannot be described categorically because tabulations by income level were not made both before and after adjustment for all the control groups. However, data on the number of families and individuals before and after the weight adjustment for 3 age of head groups—under 35, 35–54 years old, and 55 and over—as well as adjusted distributions by income level are available. Each of the 3 size distributions was given the weight of the unadjusted number it represented, and a new total distribution computed. This new total distribution was only an approximation, of course, to what the unadjusted distribution might have been. The median of the adjusted distribution was only \$12 higher than the median of the approximation to the unadjusted distribution. The proportion having incomes of \$6,000 or more was changed only 1 percent, equivalent to about two-tenths of one percentage point. This difference is so small as to afford a reason for believing that a more refined test of the 1947 weight adjustments like that for 1946 would not have revealed any significant effect on the median income. The weight adjustments are justified on another ground, however. They have a significant effect on some of the totals, presumably depicting more accurately the distribution of families and individuals by the characteristics defining the control groups, such as age of head, type of family, and color.

### 8 *Substitution*

Substitution of comparable families is not the answer since they cannot be selected with much assurance of comparability. In the BLS 1941 survey substitute addresses were selected in the field for the rural sample: the nearest house in open country, the house nearest the center of the square in villages. For the urban samples of the BLS 1941 and 1944 surveys substitute addresses were selected at random from lists in the Washington offices. A comparison of the distribution of the substitutes by income with the estimated distribution of the nonreporting families in 1941 shows that the substitute families were at a much lower level.

In the 1944 Census survey when no income data were reported for a household the nearest household (geographically)

for which a schedule was obtained was substituted. A variation of this procedure was used for missing information about persons within a household (see App. B). A summary of the effect of the Census 1944 procedure indicates that the inclusion of substitutions increased units at the \$6,000 and over level about 7 percent and affected units at other levels less. The median income was increased only \$16.

Additional evidence of the effect of office substitution procedures is available from the survey of 1947 income in the District of Columbia. The NA (not ascertained) problem was solved by office substitution of data for families of the same type, primary or secondary, and color from the same section of the city (not more than 2 census tracts distant). The median incomes of the families excluding and including substitutions differed by only \$15, and the shape of the distribution was not modified significantly.

### *9 Income Questions*

We believe that the most serious cause of bias in income data from field surveys is the inaccurate reporting of income by informants. The phrasing of questions, the detail requested, and the possibility of including probe and check questions should receive much more attention. Each of the several agencies has done something in this direction.

#### *a Wage and salary income*

According to Mrs. Goldsmith's estimates, wages and salaries comprised about two-thirds of aggregate 1944 consumer money income. Underreporting of income from this source alone would seriously affect estimates of aggregate income and size distributions. The Census schedule for the 1944 survey allowed space for only 1 figure for income from this source: wages and salaries before deductions. The BLS schedule allowed space for wages and salaries both before and after deductions, as well as for 8 specified payroll deductions. BLS interviewers were instructed to enter first the income figure most readily stated by the respondent, then to ask for detail on payroll deductions. If the respondent readily reported total income from wages or salaries,

the deductions were subtracted to yield income after deductions. If the respondent could report only take-home pay, total deductions were added to yield gross income from this source.

Of the 1,700 schedules collected by the BLS in 1944, 500 were deficient in some way with respect to the reporting of income. Of these, 162, or 32 percent (about 10 percent of all schedules collected), reported wages and salaries after deductions. Even the probes on the payroll deductions were of no help in eliciting gross income from this source. About 27 percent of the deficient schedules reported only income before deductions; the others either did not report any income or the entry was defective in some other respect.

For the 162 schedules reporting only income after deductions, income before deductions was estimated in the Washington office on the basis of income tax returns and information on deductions derived from complete reports. A comparison of the distribution of these schedules by income size throws light on the extent of underreporting by respondents who might possibly report take-home pay as total income from wages and salaries if asked only the one question.

BLS 1944 Schedules Reporting Income After Deductions Only,  
Percentage Distribution by Income Class

Income:	Income Class (thousands of dollars)					Total
	Under 1	1-2	2-3	3-4	4 & over	
Before deductions, est.	4	21	28	19	28	100
After deductions	4	26	27	21	22	100

Examination of BLS 1944 schedules indicates that as many as half of all respondents reporting wages and salaries probably gave their incomes after deductions first, then either their income before deductions or the deductions from which the before deduction figure could be derived. Since schedule entries were corrected by erasure, this cannot be stated conclusively.

Not enough information about respondents' ability to report gross income from wages and salaries was recorded on the 1944 schedules to estimate the underreporting that would have occurred had the schedule not called for both income itself and

deductions. Evidently the difficulty the BLS experienced in ascertaining gross income from wages and salaries points to the possibility of some unintentional underreporting when only the before deduction figure is requested.

This possibility does not seem to have affected the Washington figures, i.e., the median earnings of primary and supplemental earners in the Census sample. The main difference between the results of the two agencies seems to be due to the larger proportion of multiple-earner families in the BLS sample estimates, not the amounts reported by individual earners. However, the preponderance of per annum salaried employees, who presumably know their base pay, discounts the Washington experience for test purposes.

*b Farm income*

The definition of net farm income in these surveys leads to unusable results. The exclusion of family living and other non-business expenses is one desirable change. Another change that should be considered is the inclusion of net inventory increases or decreases. The great complexity of these adjustments has been the reason for failing to make them. It will be useless to alter the definition, however, unless the items are listed on a detailed worksheet. Global questions on farm income are inadequate. Schedules with itemized questions about farm receipts and expenditures seem to be the sole means of getting adequate figures since this detail would be a basis for comparing data from surveys and other sources.

*c Global estimates*

Though the assumption that the accuracy of income data reported in field surveys depends upon the detail requested seems reasonable it has never really been tested. The detail ranged from the BLS questions on individual earnings and the FRB questions on income from other sources by each member of the spending unit to the Census global questions.

The Census test of its 1947 approach to global estimates does not afford any check on the value of detailed questions (see App.

B). The three approaches were through (1) wages and salaries and income from all sources for each member of the household; (2) income from all sources for the head and related persons, and the amount from all sources for each person not related to the head; and (3), which was the same as (2) except that income was classified by \$1,000 brackets.

The median income for families yielded by the approach (1) was about \$10 more than the median yielded by (2), and about \$160 more than the median yielded by (3). On the other hand, the median for individuals yielded by approach (1) was about \$200 and \$130 lower than those yielded by (2) and (3) respectively. The standard errors of sampling are so large, however, that only the \$160 difference between the family income medians from (1) and (3) is clearly significant. Perhaps the most important point suggested by this experiment, the slight effect of detail on the amount of total income reported, is one that may be inferred also from Table 3 of Part VI. The results of the 1947 2-question schedule seem to be as accurate as those of the more detailed approaches of 1944-46, for aggregate income based on the 1947 Census size distribution was approximately the same proportion of the best comparable independent estimate as the estimates based on the 1946 Census-BAE distribution. The evidence is quite mixed whether 2 questions actually elicited lower levels of income in 1947 than 11 questions in 1946. The Census distributions were significantly below those of the FRB at many points in 1947. In 1946 the 2 sets of distributions had fewer and generally smaller differences. Other factors may well have obscured a decrease in efficiency due to the switch from the detailed to the 2-question schedule. This problem cannot be solved unless there is a controlled experiment in which one agency uses both a detailed and a global-type question schedule.

#### 10 *Probes and Checks*

The larger proportion of multi-earner families and supplemental earners, as evidenced by the Washington, D.C. area results in the BLS survey, can be attributed directly to the detailed questioning of earners with respect to their income and

expenditures. The BLS enters information on each job held by an earner on a separate schedule line. Often when the interviewer revisits the family part-time employment not originally reported is recalled and occasional and supplemental earnings are discovered. It might be assumed, although there is no factual evidence, that even intentional misreporting is sometimes corrected when the informant is confronted with a summary of his reported data that is unreasonably out of balance.

For large sample field surveys the cost of collecting income data on field schedules providing for detailed entry of each possible income item, probe questions designed to stimulate the respondent's memory, and a check balance between income, savings, and expenditures is prohibitive. Yet large samples are necessary to keep the sampling variability within reasonable bounds.

We believe that the reliability of income distributions can be considerably improved by combining data from many sources. Results of large sample field studies in which only 1 or 2 questions are asked, such as the 1947 Census, or on an intermediate type of schedule such as that used by the FRB, could be supplemented by smaller sample studies in which much detail is requested.

Even a detailed schedule, however, will not yield complete income data. In 1941 the BLS and BHNHE in a nationwide survey that used an extremely detailed and probing questionnaire accounted for about 86 percent of the aggregate income of its universe, according to Mrs. Goldsmith. The FRB in 1947 accounted for about 89 percent of its universe's aggregate using an intermediate amount of detail, questioning actual income recipients, and checking for gross inconsistencies between income, expenditures, and saving, although the coverage of the components probably varied considerably; and in 1948 for 92 percent. For this reason it is desirable that field survey data be combined with data from other sources, such as income tax returns and independent estimates of various items related to aggregate income. Combination would doubtless be facilitated if field surveys whose primary objectives are not expenditure and savings data were designed to give priority to tabulations that could

be used in making adjusted size distributions of income. The recent increase in attention devoted to the need for more controlled experimentation is encouraging.

## Appendix

### A DEFINITIONS

#### 1 *Units*

A dwelling unit was defined, in general, as a room or an apartment or other group of rooms occupied or intended for occupancy as separate living quarters by a family or other group of persons living together or a person living alone. To qualify as separate living quarters under the Census definition, a dwelling unit must have separate cooking facilities, or a separate entrance and bath, or a separate entrance and be unfurnished. If a room or group of rooms did not satisfy the definition and was in a structure with 10 or more such units (more than 10 in the case of the FRB) it was considered a rooming unit. The BLS qualified a unit as separate living quarters if the occupants were responsible for its day to day care. All other units were classified as rooming units. The FRB defined the dwelling unit as a room or group of rooms forming separate living quarters with kitchen facilities. Units without kitchen facilities, except shacks, trailers, and similar living quarters, were generally excluded from the FRB samples. All 3 agencies considered small rooming houses (fewer than 10 rooms in the case of the Census and BLS; 10 or fewer in the case of the FRB) as 1 dwelling unit.

A household consisted of the entire group of persons who occupied a dwelling unit. Transient hotels, large rooming houses, institutions, etc. were referred to as quasi-households.

The term family as used by the Census and the FRB referred to a group of 2 or more persons related by blood, marriage, or adoption and residing together; all were considered members of the same family.<sup>1</sup> Thus, if the head's son and daughter-in-law were in the household, they were treated as part of the head's family. On the other hand, a lodger and his wife were considered as an additional

<sup>1</sup> In Census release Series P-S, No. 22, units of 2 or more related persons were used instead of families, and 1-person units instead of individuals not in families.

In this paper, 'family' and 'individual' signify income-receiving units.



family, not as members of the head's family. The Census always included as members of the family unmarried children away at college unless they had established a usual residence elsewhere. The FRB and BLS included only the children absent while attending college who received their main support from their families; they excluded absent students who earned their living or lived primarily from veteran education benefits.

'Individuals not in families' as used in the Census income reports for 1945, 1946, and 1947 were persons unrelated to any person in the household.<sup>1</sup> An individual not in a family may have constituted a 1-person household by himself, or he may have been part of a household including 1 or more other families or individuals, or he may have resided in a quasi-household; e.g., a widower living by himself or with 1 or more unrelated persons, a lodger not related to the head of the household or to anyone else in the household, and a servant living in an employer's household with no relatives.

The Census primary family consisted of the head and all other persons in the household related to the head by blood, marriage, or adoption. If no person in the household was related to the head, the head himself constituted a primary individual not in a family. A household could contain only one primary family or individual. Primary families and individuals were identical with families in the 1940 Census, and the number of primary families and individuals was identical with the number of households.

The FRB spending unit was a group of persons in the same family who depended upon a common or pooled income for their major expenditures. Persons unrelated to any person in the household, classified by the Census as individuals not in families, were FRB 1-person families or spending units. An individual related to the head of a household was considered as a separate spending unit if he was at least 18 years old, had a weekly income of at least \$10, and said that he kept his finances separately and contributed less than half of his income to the family. Separate spending units may consist of 1 or more persons. The FRB combined spending unit incomes into family incomes and prepared size distributions that were comparable in this respect with the published Census data.

The BLS basic definition of the economic family was a group of 2 or more persons dependent upon a common or pooled income and sharing expenses. Unrelated persons who lived together and pooled their incomes for their major expenses were counted as one economic

family. An individual who lived apart from relatives was a 1-person family or single consumer. All persons related by blood, marriage, or adoption who lived together at the end of the survey year were considered as one economic family in the surveys for 1941-45, regardless of their financial arrangements. In surveys for these years, therefore, the economic family was comparable with the Census family, except that occasionally unrelated persons were counted as economic families.

The inconsistency between the two conditions in the BLS definition, i.e., all persons dependent upon a common income and all related persons, was clarified in the 1945 instructions when the economic family was defined as a group of related persons, regardless of financial arrangements, or a group of unrelated persons who contributed to and received a large share of their support from a common fund. In 1946 the BLS redefined the economic family in a manner that altered the basic family concept. For the 1946 and subsequent surveys the economic family was a group of persons related by blood or marriage, or an unrelated group of persons who resided together and contributed to a common income or received a large part of their support from common funds. Within a household of related persons (the Census family) there may have been 2 or more groups of persons each of which depended upon a common or pooled income, or financially independent individuals, each considered as a separate economic family. However, the BLS would have counted persons related to the family that forms the nucleus of the household as single consumer or separate economic families only if income and expenditures were clearly separated.

In concept the BLS economic family appears to have approached the FRB spending unit, as indicated by the similarity in the wording of the two definitions. However, differences in the mechanics used to determine the composition of the spending unit by the FRB and the economic family by the BLS gave rise to important differences in actual classification of the same or similar units. The BLS did not divide a group of related persons into 2 or more economic families unless both income and expenditures for major consumption items were clearly separated. For example, a son whose contribution to the family was in the form of room and board payments but who may have had use of the family car and other family possessions and services would not have been considered as a separate economic family even though he may have said he was financially independent.

The FRB would have classified a son as a separate spending unit, if he was at least 18 years old, had a weekly income of \$10 or more, said he kept his finances separately, and turned over less than half of his income to the family, even though his entire support for current expenditures came from the family fund.

The BLS recorded the composition of the economic family as it was during the survey year. Thus, a man and woman who lived as single consumers during the survey year and married just before the interview were counted as 2 individuals rather than a 2-person family. If they married during the survey year, data for the entire year were taken for the wife, but for the husband only for the part of the year after the marriage. If the woman was a member of another economic family before marriage and the man lived as a single consumer, a full-year schedule was taken for him but for her only for the part of the year after the marriage. To guide interviewers the BLS has prepared instructions covering most types of change in family composition. In general, data for part-year family members (persons who join or leave families during the survey year) covered only the part of the year during which they were members of the family. The Census and, with one exception, the FRB did not attempt to reconstruct the family, but enumerated families, spending units, and individuals as they were at the time of the interview. The exception was that in the few cases when the FRB discovered that some member of the unit had died either during the income year or just before the interview his income was included in the income of the unit.

## *2 Urban-Rural Classification*

The Census surveys followed the 1940 Census classification of urban and rural areas in which all incorporated places with 2,500 or more inhabitants were urban, together with certain other areas declared urban by special rule. Areas annexed since 1940 were classified as urban; all other areas as rural. For 1944 the BLS followed the Census definition of urban.

The Census classification of rural population as farm and nonfarm was based upon residence at the time of enumeration. Rural farm was the population living on farms at the time of enumeration in areas classified as rural in 1940. Rural nonfarm was the population not living on farms at the time of enumeration in areas classified as rural in 1940. The Census tabulated also families with farm income or loss.

In the 1941 BLS-BHNHE survey, families and single consumers who operated farms and received income from the sale of farm products were classified as rural farm. Families and single consumers who lived on farms but did not receive any income from agricultural operations were classified as rural nonfarm together with families living in villages or in open country but not on farms.

Except for 1945 FRB data are presented by Census classifications of urban and rural areas. For 1945 the rural sections of the 11 largest metropolitan areas were included with urban. The FRB did not have either a farm or rural farm classification as such, but its occupational classification defined heads of households who devoted themselves to the production and sale of farm products as farm operators. Heads of households not living on farms or working part-time at other occupations were classified as farm operators if they received half or more of their net money income from the sale of farm products, excluding however, salaries received as farm managers or laborers.

### 3 *Income*

The total money income of a person is the algebraic sum of the amounts of all types received. The total income of an income-receiving unit is the algebraic sum of the amounts received by all members. Personal taxes are not deducted.

Informants were asked to report the money income received from various sources by each recipient. The details on income requested by the several agencies and by the same agency from year to year varied considerably. Differences in the income questions on survey schedules and the techniques used in obtaining data are discussed in Appendix B.

a) *Civilian wages or salaries*, defined as the money earnings received for work performed as a civilian employee during the calendar year, included wages, salaries, commissions, tips, piece-rate payments, and cash bonuses before deductions for taxes, bonds, pensions, union dues, etc. Payments in kind, such as living quarters, meals, and clothes, were excluded (primarily because of the difficulty of evaluating them). The BLS income figures were adjusted to exclude union dues and other occupational expenses.

b) *Net income from the operation of a nonfarm business or profession* was defined as net money income (gross receipts minus expenses) from a business or profession in which the person was engaged on his own account or as an unincorporated employer. Gross receipts included the value of all goods sold and services

rendered and the value of any net increase in inventory. Expenses included costs of goods purchased, rent, heat, light, power, depreciation charges, value of any net decrease in inventory, wages and salaries paid, business taxes (not personal income taxes), etc.

c) *Net income from the operation of a farm or ranch* was computed by the Census and FRB as the difference between money receipts and expenditures. In general, living expenses were not deducted. Certain expenses, however, such as taxes (not personal income taxes), interest, and insurance, which were in part incurred by the farm living quarters and not solely by the farm business, were deducted, thereby exaggerating farm expenses in some instances. Depreciation charges, changes in the value of inventory, and the value of food, etc. produced and consumed at home were excluded.

In the 1941 BLS-BHNHE study net earnings from the operation of a farm were computed as the difference between gross income and expenses, adjusted for the value of the change in inventory. Depreciation on farm buildings and machinery was estimated as a percentage of value, and the value of the change in livestock owned and in crops stored for sale was estimated by the respondent.

d) *Armed force pay, excluding dependency allotments*, comprised the total armed force pay earned during the calendar year after deductions for Class F or dependency allotments but before soldiers' deposits, Class E or voluntary allotments, or other deductions for expenses by members of the household at the time of interview (see h below). Terminal leave pay (including terminal leave bonds in 1946) and cash allowances for subsistence were included. In the 1944 BLS survey military pay received by active members of the armed forces living at home was classified as wages and salaries. The FRB included armed force pay with income from family allotments and veterans' pensions in 1945, military pay with civilian wages or salaries in 1946 and 1947, and counted terminal leave bonds as income in 1947.

e) *Net income from rents and royalties* was defined as the cash rents and royalties received from property, including farm property, minus the costs, including depreciation charges, incurred by the landlord in connection with the property. The Census in 1945 classified royalties with other income. In the BLS 1941 study expenses incurred but not paid were not deducted. The FRB rent figures covered net income except that rents received from renting part of an owner-occupied house or apartment were gross.

f) *Net income from roomers and boarders* was defined by the Census

as gross receipts minus all expenses, such as food served to boarders, laundry, share of the wages paid to a servant for cleaning, cooking, etc., and share of the rent paid for the house (or of the property taxes, depreciation, interest, and other costs, if the house was owned). The BLS in 1941 deducted food expense from gross receipts, but not the cost of housing lodgers. The 1944 and subsequent BLS studies followed the Census procedure.

The FRB accepted gross income figures from landlords with fewer than 4 roomers, but calculated net figures for landlords with 4 or more roomers.

g) *Interest, cash dividends, and income from estates and trusts* were defined as those received or drawable in cash. Interest on Series E war bonds was not included unless the bonds had been redeemed. The Census classified income from estates and trusts with other income in 1945.

h) *Both dependency and other allotments or contributions from members of the armed forces* were included unless the person who made them was residing with the family at the time of interview, or during the survey year for BLS surveys. In the latter case, all except the dependency allotments were excluded because Class E allotments or other contributions were already included in the servicemen's pay (reported in item d above). The FRB combined these items with veterans' payments.

i) *Veterans' payments* comprise mustering-out, unemployment, disability, and other similar payments.

j) *Social security, unemployment compensation, workmen's compensation, nonveteran federal, state, and local pensions and assistance, etc.*

k) The following miscellaneous types of receipts constituted *other income*.

1) Contributions for support from persons not residing as members of the family, other than contributions from members of the armed forces. The BLS recorded large lump-sum gifts or bequests of money as other money receipts of families but did not include them in size distributions.

2) Alimony.

3) Periodic payments from paid-up endowment insurance policies or annuities, or from life insurance policies of a deceased person (including payments from war insurance). Lump-sum insurance payments and lump-sum inheritances were not included.

- 4) Cash received from private relief agencies.
- 5) Pensions from private corporations (not government pensions).

#### 4 *Universe*

The universe from which the Census and FRB samples for each year were selected was the noninstitutional population of the United States. In 1944 the BLS covered only the urban noninstitutional population.

Members of the armed forces living off post were covered. Members of the armed forces and civilian personnel living on military reservations and inmates of penal and mental institutions and homes for the aged, infirm, and needy were not covered in any study.

The Census covered more families and individuals in 1946 and 1947 than in 1944 and 1945. In 1945 the following quasi-household groups, included in the 1946 and 1947 tables, were excluded:

- a) Persons living in large lodging houses, i.e., in places having 10 or more rooms or suites of rooms rented or available for rent to roomers.
- b) Residents of hotels, YMCA's, fraternity houses, etc.
- c) Residents of trailer camps, labor camps, logging camps, houseboats, ships, etc. Residents of tourist camps and trailers were covered.
- d) Resident employees and other noninmate residents of institutions.

The same types of person were covered in 1944 as in 1945 and, in addition, persons living in large lodging houses. The 1944 BLS sample included all types of person in urban areas except those living on military reservations and inmates of institutions. The FRB covered only persons living in dwelling units as defined by it, including trailers, and excluded all residents of quasi-households, institutions, and tourists camps.

## B TECHNIQUES

### 1 *Sample Design*

All samples mentioned in this paper were area as distinguished from quota samples, and the households interviewed were selected systematically, leaving no possibility for decisions by the interviewer.

The Census estimates of income were based on data collected in connection with the Current Population Surveys. The income surveys covered about 6,700 households in 1944, about 8,700 in 1945, and about 25,000 in 1947. Data from about half the 1947 households, 12,000, were used in the published estimates. The remaining half

were divided into 2 equal subsamples for experimentation with the 2 methods of obtaining global estimates of family income. The sample households were in 68 areas in 42 states and the District of Columbia, each area comprising 1 or more counties or parts of counties.<sup>2</sup> The 1946 sample, on which the published figures are based, included about 22,000 urban and rural nonfarm households in 148 sample areas in 44 states and the District of Columbia. A few additional 1946 schedules were collected by the Census from rural farm families in the same areas, and about 3,000 schedules covering the income of farm operator families by the BAE from approximately 800 sample areas. No rural farm estimates were published, however, for reasons discussed in Section B.

The BLS income estimates were based on data obtained in connection with family expenditure surveys. The 1941 Survey of Family Spending and Saving in Wartime included about 3,000 families and single consumers. Three samples were drawn: about 1,300 in cities, 1,000 in rural nonfarm areas, and 760 on farms. The sample of urban families was drawn from 62 cities scattered throughout the country. Cities were selected by stratified sampling to give representation to 6 city-size and 8 regional classifications, racial composition, and 6 urban rent-level classifications. The number of consumer units to be interviewed in each city was based on the number of occupied dwelling units in each stratum represented, as shown by the 1940 Census. Families to be interviewed were selected at random from sample blocks stratified by average block rent in 1940. The rural samples were drawn from 45 counties selected by stratified sampling with controls on rural population, farm value, principal type of farming, state, and in the southern states, the percentage of Negroes. Within the counties the dwelling was the sampling unit, and the number of dwellings allocated to each county was in proportion to the number of occupied dwellings. The dwellings to be surveyed in the open country were taken from mile-square randomly selected areas. Villages in each county were classified by population groups, and 1 village was drawn at random from each group. Dwellings in the selected villages were chosen at random from lists or maps of all households.

The 1944 Survey of Prices Paid by Consumers covered about 1,700 families and single consumers in 28 metropolitan districts and 20 cities with populations under 50,000 selected to represent all

<sup>2</sup> The sample design is described and explained in *A New Sample of the Population*, by M. H. Hansen and W. N. Hurwitz (Bureau of the Census, Sept. 1944).



urban areas in the United States. A total of 1,400 sample dwelling units was distributed among urban places within metropolitan districts and other small cities in proportion to their population. For each urban place a sample of blocks was selected at random and all residences in each sample block were listed. From these lists residences were selected at random (see BLS *Bulletin 838*, *Wartime Food Purchases*, App.).

In the BLS 1941 and 1944 studies substitutions were made when occupants of an assigned residence refused to give the information requested or were not at home on repeated visits. The Washington office selected substitute addresses at random. In general, house-keeping units were substituted for housekeeping units, and rooming units for rooming units. In later studies the BLS selected larger samples to allow for interview losses.

In 1941 and 1944 each economic family and single consumer unit in a selected dwelling was interviewed.

The FRB income estimates were based on data obtained in its Survey of Consumer Finances; in 1945 the title was the National Survey of Liquid Assets. Income information was furnished by approximately 2,600 families and individuals not in families for 1945, 2,700 for 1946, and 3,200 for 1947. The samples were taken at 66 sampling points including the 12 largest metropolitan areas (in 1945 the 11 largest) plus 54 counties randomly selected from the remaining counties stratified according to their degree of urbanization, average per capita sales of government bonds in 1943, percentage of industrialization, percentage of native white population, average size of farm in 1940; to increase sampling efficiency by reducing sampling variance, each county in a stratum was given a chance of selection proportionate to its adult population in 1940.

Within each metropolitan area and county chosen, dwelling units were selected systematically. Each family and individual in a selected dwelling was interviewed. In 1945 the blocks within the 26 major cities that fell within the sample, representing about 40 percent of the nation's population, were stratified on the basis of the average 1940 rental value of each block. Blocks in the higher rent strata of each city were oversampled at 2-5 times the basic sampling rate of each city. Each such interview was, of course, given a reduced weight to compensate for the oversampling.

The 1946 and 1947 surveys carried the stratification within cities of over 50,000 one step further and applied it also to cities, towns, and villages with fewer than 50,000 inhabitants, which contained

about 40 percent of the nation's population. The further stratification within cities of over 50,000 required the interviewer to rate the probable income level of the families residing in dwellings that fell on certain lines of his lists. The three ratings were based upon the appearance of the dwelling unit: high, \$5,000 and over; medium, \$2,000–5,000; and low, under \$2,000. Dwellings on low rent blocks were not rated. Three strata were thus created for each set of medium and for each set of high rent blocks. All dwelling units rated high in probable income level were interviewed, and a subsample from each of the other 2 strata was selected systematically.

Dwelling units falling within a random sample of the smaller cities, towns, and villages were also rated high (a probable income of \$4,000 or more), medium, or low in probable income level. Rated units were then listed in 3 strata and subsampled at varying rates.

## *2 Income Questions*

The procedures for collecting total money income varied among the 3 agencies and from time to time in each agency's studies. Appendix Table 1 shows the sources of income for which space was provided on schedule forms.

In the 1945 and 1946 Census surveys the informant was asked to report the money income received by each person 14 or more years old, and each type was recorded separately for each person. In 1946 this was done also for persons under 14, but in 1945 incomes of children under 14 were assigned to the head of the family under other income. If any given type of income was \$10,000 or more, it was recorded as \$10,000 or more rather than as a specific amount. It was thought that insistence upon the specific amount would cause too many refusals in this income class.

In all years the FRB requested information for each member of the unit; in 1945 it recorded the data separately for the head and combined them for all other members of the unit. In 1946 it combined the incomes of all members for all items except wages and salaries. In 1947 it listed each recipient separately for all items. However, when the interviewer felt it inadvisable to insist upon detail, the FRB accepted total income figures for an earner who claimed he knew his total income. In all BLS surveys income from self-employment and from wages and salaries was entered separately for each member of the family. When more than one job was held, income from each job was recorded. Data for each other item of income were combined for all members of the family.

Appendix Table 1

Detail in which Income Data were Scheduled in Field Surveys, 1941-1947

Source of Income	BLS		Census				FRB		
	1941	1944	1944	1945	1946	1947	1945	1946	1947
Money income from all sources									
1 Wages from all sources		x <sup>a, b</sup>	x	x	x	x	x	x <sup>a, o</sup>	x <sup>o</sup>
a Earnings	x								
b Salaries	x								
c Work relief	x								
2 Self-employment		x							
a Profession	{ x <sup>d</sup>		{ x	{ x	{ x		{ x	{ x <sup>e</sup>	{ x <sup>o</sup>
b Business nonfarm	{ x <sup>d</sup>		{ x	{ x	{ x		{ x	{ x <sup>e</sup>	{ x <sup>o</sup>
c Farm	x		x	x	x		x	x	x
3 Roomers & boarders	x	x					f	x	x
4 Rents & royalties from property	x	x	{ x <sup>g</sup>	{ x <sup>g</sup>	{ x <sup>g</sup>				
5 Interest & dividends	x		x	x	x <sup>h</sup>		{ x	{ x	{ x
a Interest		x							
b Dividends		x							
6 Profits from business owned but not operated by family	x	x							
7 Military service			x				x	x	*
a Armed force pay				x	x				
b Veterans' payments		x		x	x				x
c Dependency allotments		x		x	x				
8 Social security, etc.							f	f	
a Unemp. ins. benefits	x	x							
b OASI benefits	x	{ x	{ x	{ x	{ x <sup>i</sup>				
c Other retirement benefits	x								
9 Other income	x			x	x		x	x	x
a Contributions from persons not in family	x	x							
b Direct relief	x	x							
c Periodic payment from ins., annuities, trust funds	x	x			h				x <sup>j</sup>
d Alimony, car pools, etc.		x							
e Pensions from private corp., etc.									
10 Other money receipts <sup>k</sup>		x					x	x	x

<sup>a</sup> Armed force pay included with wages.

<sup>b</sup> Earnings from each job recorded separately.

<sup>o</sup> Bonuses, overtime, commissions, and income from work outside regular job were recorded separately if not included.

<sup>d</sup> Work sheet for calculating income from business used when necessary.

<sup>e</sup> Detail recorded on withdrawals as salary and as profit and also as profits not withdrawn.

<sup>f</sup> Included in 'other income'.

<sup>g</sup> Royalties classified with 'other income'.

<sup>h</sup> State and local direct relief, including Social Security Act gratuities such as old age assistance, were classified with 'other income' in 1944 and 1945 and 'social security' in 1946.

<sup>i</sup> Income from estates and trusts classified with interest and dividends.

<sup>j</sup> Income from trust fund recorded separately on schedule.

<sup>k</sup> Inheritances, lump-sum insurance, etc. for information only; not included in total income of each unit.

Unlike the Census surveys for preceding years (see App. Table 1), the published Census income estimates for 1947 were based on schedules containing only 2 income questions for each person in the household: How much did . . . earn in wages and salaries in 1947 before any deductions were made? How much income from all sources did . . . receive in 1947?

It was felt that the first question would not only provide for comparability with 1939 but also enhance the accuracy of answers to the total income question, serving as a reminder that total income includes more than wages and salaries, a fact some respondents tend

to forget. The questions were reduced from 11 to 2 for reasons of economy and the desirability of experimenting with different sets of income questions proposed for the 1950 census.

Two other methods were tried, but the data were not published in the regular report on income in 1947. Each method attempted to get global estimates of family income and was used in half as large a sample as that which used 2 questions and on which the published figures were based. The questions in both approaches were: How much income from all sources did the head and all persons related to the head receive in 1947? How much income did each person not related to the head receive in 1947? The first was designed to yield an income figure for the entire primary family or the primary individual; the second, to yield an income figure for each other person, if any, in the household, i.e., each lodger, lodger's wife, hired hand, etc. In half the cases where these questions were used the interviewer was instructed to enter the amount on the schedule; in the other half, the interviewer displayed a card on which thousand-dollar income brackets were printed and requested the respondent to indicate the appropriate bracket; the interviewer would then enter the code for the bracket on the schedule.

The last two methods were assumed to be less expensive than the approach using 2 questions for each person. They were presumed also to be less accurate, but it was not known how much, if at all. The third method was thought to be even less expensive than the second and also to antagonize the respondent less. It had the obvious disadvantage, however, of rendering the computation of family income inaccurate for lodger and other nonprimary families. To those who were not members of primary families the question was formulated in terms of persons; otherwise, the interviewer would have had to be familiar with Census family concepts and it was felt that the burden on the interviewer was already heavy enough.

The BLS schedule for 1941 asked essentially the same question as the Census 1944-45 schedules but in more detail. In 1941 entries were made under 'earnings' for wage earners; salaried employees; and WPA and NYA earners, and unemployment, OASI, and other retirement benefits. Under the Census category 'other income' entries were made for profits from businesses owned but not operated by the family; contributions from persons not in the family; direct relief payments, and income from annuities, etc. The 1944 schedule

allotted space for income from the following sources combined by the Census: roomers and boarders, rents and royalties; interest and dividends; veterans' payments and dependency allotments (armed force pay was included with wages and salaries); unemployment insurance and retirement benefits; and profits from nonfamily-operated business, contributions from persons not in family, direct relief, periodic payments from insurance, etc., alimony and other miscellaneous recurring income, and lump-sum payments through inheritance, etc. After 1944 the BLS recorded the amount of federal tax refunded from the preceding year's payment and deducted it from the tax payment in the survey year on the assumption that the overpayment of taxes was thereby balanced.

In the BLS 1944 and all subsequent surveys the schedule had spaces for income from wages and salaries both before and after deductions. The deductions were federal income tax; OASI, unemployment compensation, railroad retirement, other retirement; war bonds; Community Chest, Red Cross, welfare organizations; union dues and assessments; group hospitalization; group life insurance; and other.

### *3 Collection and Editing Procedures*

Interviewing in each of the Census 68 sample areas (148 for 1946) was under the direction of a full-time district supervisor who was responsible for all field surveys, not just income surveys, and who recruited the field interviewers and local office assistants. The interviewers were paid by the hour at a rate varying with the size of the place. The majority were women. The income survey interviewers had usually had considerable experience in interviewing for the Monthly Report on the Labor Force or miscellaneous Census housing and business surveys. Many had also done part-time work for market research agencies. In the 1945, 1946, and 1947 surveys many had had experience in at least one preceding income survey.

In 1944 the income questions were extra items on the regular Monthly Report on the Labor Force schedule; in 1945 they were on a supplementary schedule, and the written instructions on income were more extensive. The 1946 Census income survey was part of a large comprehensive survey of the population, labor force, and housing and, as in 1945, the income questions were on a separate schedule. The extensiveness of the 1946 survey (the field work was done in April 1947) and the funds available permitted much more intensive

training for income enumeration than had previously been possible. The usual technical enumerators' instructions were supplemented by training guides and elaborate examples. Furthermore, and probably most important, a relatively large number of the Washington staff conducted training sessions with regional and district supervisors in various regional offices. The district supervisors then conducted their usual training sessions for the enumerators. One to two days were devoted to the income questions. The training procedures for the 1947 questions were the same as for 1946, but income questions were emphasized less. The income questions were the last two (in half the cases, one) on the population and labor force schedule.

Until 1947 the Census income schedules followed what is usually referred to as the record form: the entry space was designated by a simple descriptive caption and the exact questions the interviewer was supposed to ask were not printed on the schedule, as is done with the interview form. For example, on the 1946 schedule the entry space for the civilian wages or salaries was designated by a major heading, civilian wages or salaries, and a minor heading, earnings before deductions. The present Census approach is that, as far as possible, the enumerator is instructed to ask the questions as they are printed on the schedule. If the respondent does not understand the question, the enumerator is expected to rephrase it. Washington staff members on field observation trips in connection with labor force surveys had noted that enumerators did not always first phrase the questions as specified by the schedule. Whenever the phraseology was the reason, an attempt was made to improve it on subsequent schedules. Experience has proved, however, that precise adherence to carefully designed simple questions is not enough. The interviewer must be intelligent enough to know when the respondent misunderstands the question. For example, one of the present writers observed an instance in which the respondent gave a gross receipt from self-employment in response to a quite correctly phrased question on wages and salaries. Thus, the best designed schedule or instructions will not suffice unless the interviewer understands the income and employment concepts. In addition, an elementary knowledge of federal and local income tax returns has proved highly desirable.

Originally, the Bureau of the Census planned to ask income questions of all households in its April 1947 population, labor force, and

housing survey. The BAE planned a survey for January 1947 to gather various agricultural data of which income was an integral, but only one, part. To avoid a duplication of effort the Bureau of the Budget requested that the work of the two agencies be integrated so that the Census would not ask income questions of farm households. Accordingly, the BAE schedule was designed to include questions pertaining to the same nonfarm income items as the Census schedule. Net farm income, however, was to be computed in the office on the basis of the answers to specific questions about farm receipts and expenses. The Census method of collecting information about 1946 farm income from the few farm operators in its income sample was less refined (see below). It consisted of two entries, the operator's estimates of his gross and net farm income. Gross farm income was requested solely to ensure that the net figure was not in reality the operator's estimate of his gross farm income.

The forms were so different that data had to be transcribed from BAE to Census schedules at considerable cost. The BAE was of the generally more desirable interview rather than the record type, but included more questions and required a somewhat longer interview.

As originally planned, the BAE schedule included questions on such subjects as farm accidents in addition to the detailed questions on farm income. The nonfarm income questions added 6 pages. Furthermore, in some respects the arrangement of the income questions proved rather complicated. For example, information about the farm operator and about other members of the household was in separate sections; the latter consisted of 4 vertical columns for the responses of other household members who had received some income. This arrangement may have been responsible for some incomplete entries in the section devoted to the nonfarm income of other family members and of hired help and lodgers. Also, in a few cases the interviewer had evidently misunderstood the arrangement and had repeated the entries for the farm operator in the section devoted to other members of the household. The failure to follow the common practice of not making an entry, e.g., a dash, none, or a zero when no amount was reported for a specified type of income, may have been responsible for some of the ambiguity occasionally observed concerning whether a given question had been asked and whether the answer was 'None' or 'Don't know'.

The BAE used the same master sample as the Census. The seg-

ments, groups of farms, actually selected, however, were different, with possibly a few rare exceptions, and were spread over more than 800 sample areas. (As the Census sample was spread over 68 areas, the results were difficult to process.) A large number of sample areas has the advantage of reducing the between-area variance, which is relatively large for farm households. But it has two disadvantages: the cost per schedule is high and the control of field operations from Washington becomes difficult.

Because of the difficulties inherent in applying the Census definition of farm to small farms and because of problems connected with the identification as farm operators of persons who do not live on farms, it was decided that the Census survey should overlap the BAE survey by including the households of farm operators (a) having fewer than 10 acres, (b) not living on the farms they operated.

The BAE schedules representing the overlap were not machine tabulated by the Census Bureau but were hand tallied for the purpose of comparing their characteristics with those of Census schedules purporting to cover the same population group.

The practice of sending interviewers only to addresses believed to be of farms may have affected BAE results significantly. The Census sent interviewers to every address in the designated segment and they decided who was eligible to be questioned regarding income.

In all 4 years the Census schedules were edited for completeness, consistency, and reasonableness by supervisors or assistants in the field. Errors could be corrected much more readily at this stage than later, the interviewer usually being able to supply the correct information or to get it on another visit. The schedules were edited also in the Washington office.

In 1944 if a schedule lacked an entry for one or more income items, the Washington office completed the entries by the following substitution procedure before punch cards were prepared:

- a) When all income items were NA but an entry indicated that full-time, part-time, or no civilian work had been performed during 1944, the income entries for the nearest, in terms of schedule number, person with the same sex, relationship (head, wife, son, lodger, etc.), and extent of civilian work were substituted. Specifying nearest schedule number roughly introduced geographical proximity as a control.
- b) If the extent of civilian work as well as all the income items were



NA, both entries for the nearest person with the same sex and relationship were substituted.

c) If only some income items were NA, the entry substituted was that for the nearest person with the same sex and relationship for whom there was an entry of \$1 or more for the same type of income, on the assumption that when one item is reported, the informant will enter a zero for other income if none was received.

No substitutions were made in 1945 or 1946. In 1945 families with one or more members failing to report one or more items of income were entered as not reporting total money income. They were included in one of the income classes, however, in a table where the stub was a specific type of income, such as civilian money wages or salaries, on which all members of the family had reported; for example, if a respondent could give all income information except her son's armed force pay or her husband's interest and dividend receipts. Similarly, a person reporting civilian wages and salaries but not some other income item, which he may or may not have received, would be shown for some level of civilian wages or salaries, and also for some level of civilian money earnings if he had no earnings from farm or nonfarm self-employment, but would be on the persons not reporting line of the total money income level table (Census release, Series P-S, No. 22, Table 4). In 1945 no information on income or household composition was ascertained for a group of schedules representing slightly more than half a million households. Inasmuch as neither the number nor the composition of families in a given household was known, the schedules were not tabulated.

Inclusion of incomplete schedules introduced problems of presentation not ordinarily encountered. For example, if the entries for the earnings questions for a family member were all NA or were NA for one item and zero for the other items, the person's status as an earner was indeterminate; consequently, the number of earners in the family and the relationship of the principal civilian earner could not be determined. This and similar complications often made it impossible to indicate the actual number of cases represented by a given classification; for example, of 3-earner families. It likewise made it impossible to indicate the proportion of 3-earner families that did not report total family income although this information was available for some other classifications; for example, the number of 3-person families among the cards tabulated. For this reason the

not reporting cases were entered on a separate line on the simple tables in the first 1945 publications, but in the second and more detailed publication they were shown in separate tables designed to present the known characteristics of the not reporting cases.

The difficulties attending the 1945 treatment of the not reporting problem together with the dissatisfaction over the case substitution used in 1944 by the Census Bureau and previously by some other statistical agencies led to a different treatment of the 1946 data. All families and individuals whose income information was incomplete were simply excluded from the tabulations. By excluding non-reporters and adopting new and more refined inflation procedures the number of cases represented by each 1946 distribution could be shown.

NA's in 1947 were in general treated in the same way as in 1945. A major exception was dictated by the fact that 1947 income data for a given family or individual were punched on the same card as the population and labor force data. As in the case of population and labor force data in preceding surveys, whenever information concerning the composition of a certain household could not be elicited, another household of the same color in the same sample and urban-rural area was substituted. The income data in the preceding 3 years had been too detailed to be punched on the card with population and labor force data.

In 1944 and 1945 cards for persons were punched, and family punch cards were summarized mechanically. As mechanical summarization proved excessively intricate and time-consuming, the Census adopted hand summarization for the 1946 and 1947 tables.

Income data have always been a subordinate objective of BLS family expenditure surveys. On the 1941 schedule the income questions were placed at the beginning, but in subsequent surveys, because of the emphasis on price information, they were transferred to the end. In all BLS surveys field operations were directed from the Washington office. Traveling supervisors, trained by the Washington staff, hired and trained interviewers in the sample areas and generally supervised the 1941 surveys. Though by 1944 the BLS had established regional offices, the Washington staff continued to maintain direct contact with the interviewers. In these two studies schedules were reviewed chiefly by the technical departmental staff; sometimes follow-up on deficient schedules was impossible. Since 1944, when

the surveys were confined to 3 selected metropolitan areas a year, management problems have been simpler. Regular BLS employees were the supervisors. After extensive training in Washington by members of the technical staff in charge of the studies, they established offices in the survey areas, recruited and trained interviewers, and directed the day to day operation of the study and the editing and review of completed schedules. Field editing and review permitted almost immediate follow-up by the interviewer to get missing data and to correct inconsistencies in entries.

The chief criterion for determining the goodness of the data was the schedule balance. If the sum of expenditures and change in unit net worth deviated more than 10 percent from the reported unit total money income, the interviewer attempted to reconcile the reported data by further personal contact with the reporting unit. Usually reported income was lower than expenditures. The interviewer, on a revisit, tried first to discover unreported income, then errors in savings and expenditures.

After the schedule had been submitted to the Washington office, the income section was again reviewed for completeness and consistency. The following items were checked:

- a) In the family composition section the interviewer indicated by check mark each family member employed, and explained the status of any person over 18 not employed 13 or more weeks during the year. This record was checked against the record of earners in the income section of the schedule.
- b) Income from self-employment was checked to determine whether all business or occupational expenses (such as business use of the automobile) had been deducted.
- c) Wage and salary income was checked by comparing income before deductions with the sum of reported income after deductions and the amount deducted. Deductions were reviewed for completeness and such deductions as taxes, OASI, and unemployment compensation were estimated if not reported. Deductions for the purchase of savings bonds were checked against reported bond purchases. If either wage and salary income before or after deductions was not reported, the missing item was estimated.
- d) Other income items were subjected to only a minimum of editing unless the interviewer indicated on the schedule that a substantial amount could not be reported by the informant.

The BLS tried to reconstruct the family as it existed in the survey year, and included in total family income only the income each individual in the reconstructed family received while a member of the family. In this way it corrected in part for the error that occurs when a sample is drawn at one point in time to represent a population as it existed earlier. Reconstructing the family is difficult, presenting chances for errors that may be more serious than the biases it attempts to correct. An examination of field instructions issued to agents since 1944 indicates that the BLS did not follow a complete reconstruction procedure. This may be attributed to expediencies resorted to for simplifying in some degree the rather difficult expenditure surveys.

When the family changes in composition between the end of the survey year and the interview, the problem is not complicated by the need for including information about part-year members. The agent determined the composition of the family at the end of the year and took full-year data for each member. To avoid possible double counting, BLS agents were instructed not to take schedules for "families consisting entirely of persons who were members of other economic families at the end of the year". Thus a man and woman who were members of other families during the survey year and married after the year-end would not be scheduled if they happened to be selected in the sample; their incomes would be included in the total incomes of the families they left, if those families were selected.

In accordance with the rule for reconstruction no schedule was taken for either a man who married after the year-end or his wife if either or both lived apart from relatives during the entire survey year; and no schedule was taken for the independent man or woman who married into the spouse's family after the year-end. For complete reconstruction, these 'individuals not in families' should have been scheduled individually. Their exclusion tended to bias the income distribution upward, since a high proportion of the excluded individuals were at the lower income levels.

The problem of avoiding double counting by reconstructing the family requires extreme care in writing instructions and training agents. The BLS rules covering marriages during the survey year illustrate this point:

"(1) If both persons were separate economic families at the begin-

ning of the year, a full year schedule is to be taken covering the woman from the beginning of the year and treating the man as a part-year member.

(2) If the woman was a member of another economic family and the man was a separate economic family, a full year schedule is to be taken covering the man from the beginning of the year and treating the woman as a part-year member.

(3) If a man or woman marries into a family of 2 or more persons, a full year schedule is to be taken covering the family and treating the additional member as a part-year member."

These rules helped field agents to recognize family situations easily. Despite all precautions, BLS instructions still contained loopholes that allowed some duplication. For example, a schedule could be taken for both husband and wife from either divorced member of a broken family; or complete data for an undoubled family might be taken from either of the new families after the undoubling. But the probability of drawing two groups from which duplicate schedules could be taken is small.

The FRB has always collected income information in its general surveys of the financial situation and expectations of consumers. The interviews were begun with fairly long explanations of the nature of the survey, followed by a few general questions designed to arouse the interest and cooperation of the respondent. The income questions were then asked. All questions had been pretested for clarity and understandability and their phrasing and order had been reworked until they were a smoothly flowing, easily understood set. The interviewer was required to ask the questions exactly as printed.

The average interview took about an hour. In covering many financial matters as well as income, such as purchases of houses, automobiles, sales of stocks and bonds, previously forgotten items of income were sometimes recalled. The interview, while not eliciting detailed data on expenditures, sometimes revealed gross inconsistencies in the data on income, expenditures, and savings, which brought to mind overlooked income items.

The collection of interviews at each sample point was under the direction of a full-time supervisor from the Survey Research Center. Each supervisor had been trained for almost a week at the central office, during which the purposes and meaning of the survey questions were thoroughly explained. Beginning with the 1946 survey,

each supervisor conducted an unrehearsed interview before a recording machine; transcribed, these interviews were then criticized by the other supervisors and the central office staff, and used later by the supervisors to train the local interviewers.

The supervisors visited each sample point to hire and train interviewers. A large proportion of the local interviewers had previously worked for the Survey Research Center. Each interviewer, during the 2 to 3 days of training, made several practice interviews which were carefully analyzed by the central office staff.

The schedules were thoroughly checked for omissions and inconsistencies in the central office. When necessary, they were returned to the interviewer. Finally edited and coded, the information was transferred to cards for machine tabulation.

A considerable effort was made to hold nonresponse to a minimum. A letter indicating the nature of the survey in general terms was sent in advance to each household to be interviewed notifying it that an interviewer from the Survey Research Center would call during the next few days. When no one was at home, the interviewer attempted to arrange for an appointment. In urban areas at least 3 calls were made at different times of the day and evening; in rural areas at least 2 calls.

When the interviewer met with a refusal, he noted the circumstances on a special form and sent it to the central office. The central office then prepared an individual letter explaining the purposes of the survey and its anonymous use of the requested information, and the interviewer called again.

No substitutions were permitted, but the sample was large enough to compensate numerically for nonresponses. The response rate was quite similar for all 3 FRB surveys. Data for 1946 are presented in terms of spending unit nonresponse, since they are unavailable on a family or individual basis.

Interviews from 86 percent of the spending units in the sample were accepted. The reasons for failures in the other 14 percent were as follows: 6 percent were not at home or did not keep appointments; 1 percent were unavailable because of sickness or death in the family, language difficulties, etc.; and 7 percent refused to give information or gave such incomplete data that the schedule was rejected.

Of the accepted interviews, 99 percent gave the income details requested. One percent gave either no or incomplete income data

but gave information on liquid asset holdings, raising the total partial and complete nonresponse rate for income to 15 percent. In the first two surveys assignments for this 1 percent were based primarily upon the liquid asset holdings of the unit. Beginning with the third survey, which covered income in 1947, they were based on savings data as obtained from reported changes in various assets and liabilities of the spending unit.

A further problem of incomplete reporting arose in the process of combining related spending units into family units. A little over 3 percent of the accepted family and individual units lacked income information for a spending unit other than that including the head of the household. When information was lacking for the head of the household's spending unit (the primary spending unit), the schedule for the entire family was rejected. For the 3 percent missing related secondary spending units, information was supplied through stratifying reporting multi-spending unit families on the basis of geographic areas and sampling rates, income of primary spending unit, and ranking secondary spending units by size of income. The incomplete families were stratified in the same manner and matching cases selected to complete them by random selection within identical strata. Thus the distribution of reporting families with the stratified characteristics was maintained.

The rate of nonresponse varied markedly among strata. For groups sampled at the basic rate and at somewhat above the regular rate, total nonresponse was 14 percent, mounting to 29 percent of the spending units sampled at much more than the regular rate.

Different kinds of geographic area, sampling area, and type of unit (primary and secondary) were adjusted separately within each stratum by weighting upward reporting cases to compensate for nonreporting in the stratum.

#### *4 Inflation of Sample Data*

In 1944 and 1945 the Census income sample results were multiplied by the unadjusted weights (reciprocals of the sampling ratios). A Census Bureau estimate of the number of primary families and individuals (identical with the number of households) was then divided by the number of such families and individuals indicated by the field surveys. The weights, the number of cards to be punched for a given person or family, were multiplied by the quotient of the

foregoing division. The effect was to raise the sample weights a little more than 1 percent, the degree to which the sample estimate of the number of households fell short of the independent, and presumably more accurate, estimate of households.

The FRB followed a similar procedure for 1945, 1946, and 1947, inflating its sample results to an estimate of total dwelling units in the universe, based on data prepared by the Census Bureau for a given date.

In 1946 the Census inflated the weighted sample results to agree with independent estimates of the civilian population with respect to age, sex, color, and veteran status of males, based on the 1940 Census of Population, statistics of births, deaths, and net immigration, and the strength of the armed forces. The proportions of the population in urban, rural nonfarm, and rural farm areas, derived from the concurrent sample Survey of Population, Labor Force, and Housing, were superimposed on the controls listed above. Income survey punch cards for families, individuals, and persons were then collated so that each family card was matched with the principal person card for that family, both being given the same weight. The principal person was the wife in husband and wife families and the head of the family in other cases. The former was justified on the ground that in preceding sample surveys when the number of persons, estimated by multiplying the sample figures by unadjusted weights, generally did not equal the independent estimates, the relative deficiency was less for females than for males.

The family and individual cards were then tabulated by several characteristics and the results compared with the figures from the sample Survey of Population, Labor Force, and Housing which, with a few exceptions, covered the same households as the income survey in the urban and rural nonfarm population. Since returns for which income information was lacking or incomplete were not tabulated, discrepancies were to be expected for the urban and rural nonfarm classifications because of the bias created by the higher rate of nonreporting on income than on other schedules. Rural farm discrepancies were to be expected for the same reason but perhaps more because most of the rural farm sample was taken by the BAE and from a different source. The comparison of the income sample estimates with the control figures from the Population, Labor Force, and Housing Survey revealed an underestimate of more than



4 percent for individuals not in families, due to their higher NA rate especially among lodgers, and an overestimate of about 1 percent for families. The discrepancies were large enough to justify adjusting the weights further.

Appendix Table 2

Families and Individuals, Percentage Distribution by Total Money Income Classes, Rural Nonfarm and Farm, Census Surveys for 1944, 1945, and 1947 and Census-BAE Survey for 1946

Total Money Income Class	1944		1945		1946		1947	
	Families & indiv.	Families	Families & indiv.	Families	Families & indiv.	Families	Families & indiv.	Families
	<i>Rural Nonfarm</i>							
Loss	...	...	...	0.1	0.6	0.2	0.2	0.2
0 - \$499	12.0	7.1	9.5	4.7	9.6	4.9	8.8	4.3
500 - 999	10.3	9.5	9.6	7.4	9.8	8.0	8.4	6.6
1,000 - 1,499	12.4	11.5	9.1	8.8	10.0	10.1	8.3	7.7
1,500 - 1,999	12.1	12.6	14.8	15.4	11.2	11.6	10.2	10.5
2,000 - 2,499	11.1	11.8	14.0	15.4	12.9	14.1	12.4	13.6
2,500 - 2,999	10.4	11.2	11.8	13.2	10.5	11.7	10.3	11.0
3,000 - 3,499	10.6	12.2	9.1	10.2	9.7	10.6	11.7	13.0
3,500 - 3,999	6.2	7.0	5.6	6.2	6.7	7.5	8.3	9.1
4,000 - 4,499	3.5	4.1	5.1	5.8	4.7	5.3	5.7	6.3
4,500 - 4,999	3.4	3.9	3.3	3.7	3.5	4.0	3.5	3.9
5,000 - 5,999	4.0	4.6	3.5	4.0	4.4	4.8	5.4	6.2
6,000 - 9,999	3.6	4.1	3.8	4.2	4.7	5.3	5.3	6.0
10,000 & over	0.4	0.4	0.8	0.9	1.7	1.9	1.5	1.6
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Median inc., \$	2,145	2,388	2,248	2,445	2,330	2,548	2,581	2,826
	<i>Rural Farm</i>							
Loss	1.7	1.7	1.6	1.5	7.9	7.8	0.7	0.7
0 - \$499	22.7	19.5	21.6	17.9	17.9	15.1	14.6	11.0
500 - 999	20.7	20.0	18.9	19.0	17.3	17.1	14.7	13.4
1,000 - 1,499	15.7	16.3	13.6	14.2	12.5	12.8	13.2	13.4
1,500 - 1,999	9.7	10.4	13.8	14.4	10.0	10.4	11.9	12.5
2,000 - 2,499	9.3	10.0	10.1	11.1	8.4	9.0	9.7	10.8
2,500 - 2,999	5.6	6.1	6.5	7.1	6.3	6.7	6.9	7.5
3,000 - 3,499	4.2	4.6	4.0	4.3	4.7	4.9	6.9	7.4
3,500 - 3,999	2.9	3.2	2.1	2.3	3.1	3.3	3.1	3.4
4,000 - 4,499	1.7	1.9	1.9	2.1	2.8	3.1	3.6	3.7
4,500 - 4,999	1.2	1.3	1.9	1.9	2.1	2.3	1.7	1.9
5,000 - 5,999	1.5	1.7	1.5	1.6	2.1	2.2	4.6	5.1
6,000 - 9,999	1.6	1.7	2.0	2.0	3.5	3.8	5.7	6.3
10,000 & over	1.5	1.6	0.5	0.6	1.4	1.5	2.7	2.9
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Median inc., \$	1,157	1,272	1,291	1,410	1,275	1,390	1,781	1,958

To revise the weights in preparation for the tabulations that were to be published, income cards for persons were duplicated or eliminated to bring the figures from the two sources into agreement with respect to age, sex, race, and veteran status (males only).

For primary families and individuals the figures from the two sources were made to agree by urban-rural area with respect to the sex of the head (2 age groups) and age for male heads; for other

families by 6 population survey control groups, a classification by household or quasi-household residence within each of the 3 urban-rural categories. For other individuals there were just 2 control groups, household and quasi-household.

The 1947 weights for persons were adjusted to make the sample estimates for persons agree with independent estimates of the number by age, sex, and veteran status (of males).

Weights for males in the armed forces received the same adjustment as those for male nonveterans in comparable age groups. The family and individuals cards were then grouped according to the following family individual types:

- 1 Primary families
  - a Husband and wife
  - b Other, male head
  - c Other, female head
- 2 Primary individuals
  - a Male
  - b Female
- 3 Other families
  - a Husband and wife
  - b Other, male head
  - c Other, female head

FRB weighted sample results were inflated by means of the independent estimates of the number of dwelling units in the survey universe prepared by the Population Division of the Census.

### *Comments*

WILLIAM VICKREY, *Columbia University*

If I may be permitted to ride a hobby of mine for a bit, the excellent discussion in this paper seems to me to bring further support to the proposition that it is of primary importance, both in designing a study and in tabulating the results, to consider analytical significance as well as statistical convenience in choosing the unit to be polled and the basis upon which the results are to be tabulated.

Apparently, income (also certain minor items of savings and consumption, as well as the degree of part-time employment) can be accurately ascertained only through the use of detailed schedules, together with relatively time-consuming and thorough interview techniques. This means that if the expense of the project is to be kept within bounds, the sample must be small. But if the sample is small, the number of cells in the classification scheme must likewise be reasonably small, otherwise the cases in each cell will be too few to give a reasonably small random sampling variance. Much as we would like to have data cross-classified by income level, family size, region, type of community, occupation, age of family head, and perhaps even items reflecting past history (to mention only a few of the relevant criteria), in many if not most cases we shall have to content ourselves with classifications by a relatively small number of criteria at any one time. Consequently, the contents of each cell will necessarily be heterogeneous in some respects. The errors and possible biases introduced by this heterogeneity, however, can be substantially reduced if we take care to express the criteria for classification in ways that minimize the relevant heterogeneity when the fineness of the classification is limited to any given number of parameters, or degrees of restraint.

If, for example, we wish to study the saving-income ratio as affected by ethnic origin, income level, and size of community, we might ideally also want, as a fourth degree of restraint, the size or type of family. Owing to the small size of the sample, however (and possibly also to the time available for analysis), the more detailed classification and analysis will often have to be foregone. If we are thus restricted to less detailed analysis, we are likely to get more homogeneous groups within each cell, and hence more reliable results, if income level is determined in terms of a quotient of family income divided by some index of family size than of an unadjusted total income per family. To be sure, in principle the results would be the same if we classified further by type of family. Indeed, if time sufficed and the sample was large enough, more information could be obtained in this

way. In practice, however, since both the number in the sample and the time available for computation and tabulation are always limited the number of cells that can be set up is likewise limited.

Again the proposal that cross-checking among surveys be facilitated by adopting a common definition of the unit and classification of the data has a bearing on the methods of classification to be adopted. Of course, the common classification method must be one that requires only data obtainable from all the various types of survey. However, even if the definitions used in the several surveys are in principle identical, each survey will for one reason or another classify units differently. There will always be borderline cases different enumerators will treat differently. Families may combine or break up, or some members may leave for seasonal jobs, so that the same enumerator may classify the family differently according to when the enumeration happens to take place. Finally, the questionnaire, differing according to the purpose of the survey, may turn up different amounts of information that may affect the way the unit is classified. A suitable choice of method can considerably reduce the effect of such differences in treatment on the final results.

For example, a group of 8 persons with a total income of \$10,000 might be considered by one enumerator to be a single consuming unit, and by another to be composed of two consuming units, one of 5 persons with \$6,000, and another of 3 persons with \$4,000. If the results of the two surveys are tabulated on some basis such as 'per capita income' or 'income per equivalent adult', both surveys will rank the 8 individuals and the \$10,000 at about the same percentiles on the income scale, the only difference being the number of units reported; and the error of enumeration will cause relatively little difference in the over-all distribution of income or in the patterns of savings or consumption. On the other hand, if the basis of classification is income per family or per consumer unit, the two surveys will rank this group at substantially different percentiles on the income scale, and the distribution of income and other patterns may differ considerably.

Another finding with interesting implications is that the longer the interval between the period covered by the study and the interview, and by implication, the longer the period covered, the greater the chance for errors to creep in. We could, of course, get much better information if we inquired about income and expenditure for only the preceding week or the preceding month; seasonal and other influences would, however, impair the usefulness of such material. Or perhaps interviews repeated at frequent intervals throughout the year might permit a fairly accurate picture to be built up of the transactions for the year. However, interviews during the period reported on tend to influence the behavior of respondents, so that they in some degree will no longer fairly represent the un-interviewed portion of the community. Not only do the interviews consume considerable time, but respondents may gradually become self-conscious or budget conscious about certain expenditures and so shift their patterns of consumption and perhaps investment in a manner not representative of the group as a whole. A bias due to the influence of the interviews is more or less inevitable, however, if we try to collect data covering periods longer than a year, as will be necessary if we are to have adequate data for studying the effects of changes in income on consumer behavior, or the relation between distributions of annual income and of national resources. This problem is partly met by carrying forward both a 'panel' of respondents who are interviewed repeatedly to obtain a record of income and spending patterns for several years, and a series of annually renewed samples that would be kept free from the bias due to the influence of the interviews. But the knotty problem of dealing with families that form and break up still remains. One may perhaps detect the emergence of a 'Heisenberg principle' in the social sciences: if one tries to measure a phenomenon with more than a certain degree of accuracy one causes changes in it that in turn resist accurate measurement.

A. ROSS ECKLER, *Bureau of the Census*

The Census Bureau wishes to take this opportunity to acknowledge the very great contributions that have been made through

this study of the field surveys of income. We have learned a good deal from the work of the authors and hope that as a result we can improve our income statistics.

We are very much interested in any work aimed at evaluating the accuracy of data collected in field enumerations. We are studying the problems of response variation and its causes. To evaluate the accuracy of census and survey figures we have conducted quality checks and propose to include them in all major censuses.

So far, our quality checks do not entirely support Mr. Vickrey's contentions that more and more questions improve results. However, it is too early to attempt to reach any final conclusions concerning the merits of intensive interviewing. In some cases we have found that the addition of questions on sources of income increases the accuracy of the total income figure obtained, but it is not by any means clear that the gains are sufficient to justify the cost.

In view of the amount of work that needs to be done in this field and the present state of our knowledge, I would like to commend especially the authors' conclusion that field surveys of income are only now beginning their scientific stage and that there is "need for more controlled experimentation".

I would like to comment briefly upon the desirability of adjusting field surveys of income in order to get a better over-all national income distribution. We in the Bureau of the Census recognize that the data collected in field surveys have certain deficiencies which can be corrected for by the judicious use of statistical techniques. We believe that the staff working on these adjustments has done a competent and painstaking job and that the adjustments carried out have been made with full knowledge of the methods used in our surveys. We hope that this kind of work continues, not only because we believe it improves the income distribution series but also because we expect that the processes of adjustment will reveal certain shortcomings in the data which may enable us to design better schedules and improve our program of training and supervision.

One element that should perhaps be taken into account is the persisting difference between income data obtained from field surveys and from establishment reports. Apparently under the most favorable conditions, the gap has never been less than 7 percent and is usually 15-20 percent. It is possible, as Jerome Cornfield recently suggested, that the underenumeration of the population characteristic of all censuses and surveys may be held largely responsible when the discrepancy reaches a level as low as a few percent. We shall know better after we have completed the 1950 Census together with a quality check of the results.