

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: Estimating the Number of Earners for Income Size Distribution Analysis

Chapter Author: Emmett Welch

Chapter URL: <http://www.nber.org/chapters/c5731>

Chapter pages in book: (p. 560 - 571)

# *Part X*

## Estimating the Number

---

# of Earners for Income Size Distribution Analysis

*Emmett H. Welch*

NATIONAL SECURITY RESOURCES BOARD

---

ESTIMATING THE NUMBER OF EARNERS is a major part of analyzing information on income size distributions. Population-based income surveys ordinarily collect information on the number of earners as an essential step in obtaining information on the amount of earnings. Usually, for example, before a person is asked how much he earned, he is asked whether he had any earnings. It is generally recognized that the amount tends to be underreported. What is not as generally recognized is the extent to which the receipt of any earnings at all tends to be underreported, and the effect that this has on the size distribution of income.<sup>1</sup>

<sup>1</sup> This paper deals only with underreporting of the number of persons with earnings. The number of persons with income other than earnings is probably subject to relatively more underreporting.

#### A EVIDENCE OF UNDERREPORTING OF THE NUMBER OF EARNERS IN INCOME SURVEYS

The survey of 1944 income conducted by the Bureau of the Census in May 1945 found that 52 million persons in the non-institutional population 14 years and older had received civilian earnings. This figure is obviously low since the Monthly Report on the Labor Force recorded civilian employment, excluding unpaid family work, at about the same level for the peak month in 1944. Allowing for persons employed at some time during the year but not in the peak month would bring gross civilian employment for the year considerably above this figure.

For 1945 the number of persons reporting civilian earnings was also about 52 million while civilian employment excluding unpaid family work was about 51 million in the peak month. For 1946 the corresponding figures were 60 and 55 million. The labor force turnover was probably somewhat higher in 1946 than in 1944 or 1945, increasing the number of earners during the year relative to the peak. This would account, in part, for the wider difference in 1946 than in 1944 and 1945 between the number of earners and peak employment. In part, however, this wider difference indicates less underreporting of the number of earners in 1946.

The 52 million earners reported for 1945 in the Census income survey and the 60 million reported for 1946 were 20 percent below the 65 million for 1945, and 14 percent below the 70 million for 1946 estimated by Selma Goldsmith (Part VI). The Bureau of Old-Age and Survivors Insurance has recently estimated 71 and 73 million for 1945 and 1946, but these figures include five groups not included by either the Census Bureau or Mrs. Goldsmith: inmates of institutions, members of the armed forces, persons who died prior to the date of the survey, children under 14, and residents of Hawaii and Alaska. The total number in these categories who have civilian earnings is not likely to exceed 2-2.5 million. Reducing the BOASI estimates by this number would leave them higher than those prepared by Mrs. Goldsmith. Both the BOASI and Mrs. Goldsmith's estimates

are based on various bodies of data, primarily the number of persons with wage credits under the Social Security Act. These estimates indicate not only that the number of earners reported in the Census income surveys is too low, but also the possible variation in estimates based on similar data.

Additional evidence of underreporting is provided by experiments recently undertaken by the Bureau of the Census in preparing for the 1950 census. In May 1948 a test census with various questions being considered for the 1950 census was taken of Cape Girardeau and Perry Counties, Missouri. Questions on earnings and income were included. As part of the test, members of the Washington Census Bureau staff, together with selected enumerators and supervisors who had worked on the original enumeration, re-interviewed about 1,000 households with the object of verifying the original enumeration. Some 5 percent of the persons reporting no wages or salary in the original interview reported them in the re-interview. About 2 percent of those who reported no wages or salary in the re-interview had reported them in the original interview.

Some of these differences can be ascribed to the variation in response that always occurs because of differences in enumerators, respondents, and timing of the field work. Some of those who reported in one interview but not in the other that they had received income probably had not received any income. To a considerable extent, however, the differences are likely to represent underreporting of the number of persons with wages or salaries. The fact that the re-interviewers failed to obtain information about wages and salaries reported in the original enumeration suggests that they identified only some of the wage or salary earners who had been reported as nonearners in the first enumeration. It may be inferred, therefore, that this test indicates that more than 5 percent of those reporting no wages or salary had, in fact, received wages or salary during the calendar year preceding the survey. Since about half of the persons 14 and over had reported wages or salary, the number of earners was underreported at least 5 percent.

The sample in this pre-test was very small and concentrated

geographically. It is not desirable, therefore, to attempt to deduce the general rate of underreporting in surveys of this type. Nevertheless, the 5 percent seems small when compared with the 20 percent difference between Mrs. Goldsmith's and the Census income survey estimates of the number of earners in 1945 and the 14 percent difference in 1946. The difference between the Census and the BOASI estimates, adjusted for coverage, was 24 percent in 1945 and 15 percent in 1946.

#### B EMPLOYMENT QUESTIONS IN THE CURRENT POPULATION SURVEY

In connection with the Monthly Current Population Survey the Bureau of the Census has, on several occasions, asked questions concerning work done during the preceding year. From the answers the number of earners during a year can be estimated. In August 1946, for example, persons not in the labor force were asked whether they had ever had a job lasting two weeks or more. If they answered yes, the date the latest job ended and its type were inquired. This made it possible to estimate the number of persons currently not in the labor force who had had earnings from a job lasting two weeks or longer during the preceding twelve months. About 7 million persons not in the labor force during the survey week were estimated to have had earnings from such jobs. Adding these 7 million to the 55 million reported as earners during the survey week gives 62 million. To arrive at the total number who had earnings during the preceding twelve months it is necessary to add the estimated number of unpaid family workers and unemployed persons in August who had been earners at some time during the preceding twelve months.

In the December 1947 Current Population Survey about 20 percent of the unpaid family workers were reported as having been earners at some time during the year. The proportion would be somewhat lower in August because of the larger proportion of students temporarily working as unpaid family workers during the summer vacation. It may be estimated, there-

fore, that about 500,000 of the 2,800,000 unpaid family workers in August 1946 worked for pay or profit during the preceding twelve months.

Some 850,000 of the 2 million unemployed in August 1946 were recently discharged veterans who, by and large, had not found their first civilian job since leaving the armed forces. Of the remaining 1,200,000, a good many were young people looking for their first job. In the December 1947 survey about an eighth of the unemployed had not been earners at any time during the year. In August, when the ranks of the unemployed are swelled by students on vacation, this proportion would probably be higher. Apparently, perhaps a million of the unemployed in August 1946 had been earners at some time during the preceding twelve months.

In summary, therefore, we have 63.5 million persons 14 and older in the civilian noninstitutional population in August 1946 who had been earners at some time during the preceding twelve months. This is not a complete count of all earners, since it does not include those not in the labor force in August 1946 who had had earnings from a job lasting less than two weeks. Also, there is considerable evidence that some respondents who supply information to the Census Bureau for other persons in the household do not report the work experience of some individuals, either because they do not know about it or have forgotten.

Persons with earnings during the survey week	55,000,000
Persons not currently in the labor force who had been earners	7,000,000
Unpaid family workers who had been earners	500,000
Unemployed who had been earners	1,000,000
Total	63,500,000

Converted to a calendar year basis, the 63.5 million for the 12 months ending August 1946 would become 65-66 million earners for 1946, considerably more than the Census survey figure, 60 million earners, but less than Mrs. Goldsmith's estimate, 70 million.

Even when a series of questions on work during the year is asked the number of earners seems to be underreported. For example, in December 1947, respondents were asked whether

each person 14 and over had during 1947 (a) done any farm work for cash wages; (b) done any nonfarm work for wages or salary; (c) operated a farm; or (d) operated a business or worked at a profession.

According to this survey, about 64 million persons 14 and older in the civilian noninstitutional population in December 1947 did some work for pay or profit in 1947. In publishing the results, the Census Bureau indicated that the number of earners was probably higher (*Work Experience of the Population in 1947*, Series P-50, No. 8, Aug. 13, 1948):

"It is probable that the figures presented in this report understate the number of persons who did any work for pay or profit in 1947. Occasionally, respondents, either because of incomplete knowledge or because of inability to recall, did not report the paid work experience of some individuals. Students on vacation from school and many housewives, for example, are outside the labor force most of the year, and respondents, in some cases, may have been unaware of or unable to recall the fact that such persons had in fact been employed several months earlier."

About 80 percent of the 25,000 sample households enumerated had been in the sample also for four months prior to December. For them a record of the responses for each month August–November was available. A check of these records showed that about 1,100 or 5.5 percent of the 20,510 persons in these sample households who were reported as not having done any work for pay or profit in 1947 had, in fact, been employed as wage or salary workers or had been self-employed in at least one month. In addition, there must have been others who did such work between January and July but not between August and November. An estimate of the error in the rate of response in the December survey should allow for the additional errors that would be uncovered if the records of previous work experience covered all eleven months prior to December instead of only the last four.

The records for the four months indicate to some extent whether these additional errors would be few or many. If the



records for persons erroneously reported as not having worked for pay or profit showed that these persons did such work for several months, it could be inferred that such work experience tended to repeat itself, so that persons who worked in the first part of the year probably worked in the last part also; few additional errors would therefore be uncovered if the records of previous work experience could be extended back to January. If, on the other hand, the records showed that the paid employment was an isolated experience which occurred only in a single month, one could conclude that many additional errors would probably be disclosed by extending the work experience records back to January.

Of the persons reported as not having worked for pay or profit in 1947 who did work for wages or salaries or as self-employed between August and November 64 percent worked in one of the months; 19 percent in two of the months; 10 percent in three of the months; and 7 percent in all four months.

As apparently most persons who were erroneously reported as not having worked for pay or profit did such work in only one month, it may be inferred that the records for August–November disclosed only a portion of those who worked for pay or profit in 1947. Of those reported as not having worked for pay or profit in 1947 how many did such work between January and July in addition to the number already disclosed as having done such work between August and November? The table shows roughly the effect of adding more months to the record of previous work experience checked to discover reporting errors.

Over 400 persons among those erroneously reported as not having worked for pay or profit in 1947 worked in August but not in any subsequent month. This figure is unusually high because there is more seasonal and short-term employment in August than in any other month of the year. Over half of the persons 14–19 known to have been erroneously reported as not having worked for pay or profit in 1947 worked in August but not in any subsequent month. Among persons 20 and over, this proportion is less than one-third. The particularly high pro-

portion for August among those under 20 is due to the fact that many youths, usually occupied as students, take temporary jobs during summer vacations.

Persons Reported as not Having Worked for Pay or Profit in 1947 Who Worked for Wages or Salaries or as Self-Employed August–November by Most Recent Month in Which They Worked

	Total	Most Recent Month in Which They Worked			
		Nov.	Oct.	Sept.	Aug.
		<i>All persons 14 and older</i>			
Number	1,092	262	220	207	403
Percentage	100	24	20	19	37
		<i>Persons 20 and older</i>			
Number	741	209	161	151	220
Percentage	100	28	22	20	30
		<i>Persons 14–19</i>			
Number	351	53	59	56	183
Percentage	100	15	17	16	52

To estimate the number that more complete records would disclose as having worked in July (and other months back to January) but not in any subsequent month, we must remember that

- 1) The difficulty in recalling work experience will probably be greater the longer the interval between the date of work and the date of interview. This factor would tend progressively to increase the figure for July and each prior month.
- 2) The number of persons who worked in a given month but not in a subsequent month would tend to be less as the number of subsequent months increased. This factor would tend progressively to decrease the figure for July and each prior month.

The table suggests that the latter factor tends to outweigh the former slightly. Excluding the figures for August, which are affected by seasonal factors, the figures for earlier months decrease progressively for all persons 14 and older and for those 20 and older. For persons 14–19 there does not appear to be any significant trend.

Although an estimate of those erroneously reported as not having worked for pay or profit in 1947 who last worked each month January–July is difficult with the limited data at hand,

an average of about 120 for those 20 and older and about 50 for those 14-19 would not seem unreasonable. Making this allowance for the months prior to August would add 1,200 to the 1,100 already shown, bringing the rate of underreporting of earners to 11 percent of those reported as not having worked for pay or profit in 1947.

Some adjustment should be made also for the fact that the records against which the December 1947 survey was checked covered only one week in each month. Data are not available upon which to base a reasonable estimate of the degree to which the rate of underreporting would be raised if all weeks could be checked.<sup>2</sup>

According to the December Current Population Survey, about 42.5 million persons were in the civilian noninstitutional population that month who did not work for pay or profit in 1947. On the basis of the evidence adduced in the preceding sections, it seems that at least 11 percent, about 5 million, did, however, work for pay or profit. Adding 5 million to the survey estimate of the number who worked for pay or profit in 1947, 64 million, gives a revised estimate of 69 million.

Probably an estimate of the number of earners for 1947 made with Mrs. Goldsmith's methods would not be much different from her estimate, 70 million, for 1946. It appears probable, therefore, that if the Census Bureau estimate, 69 million, were adjusted to allow for weeks that could not be checked for underreporting, the two estimates would approximately agree.

### C YEARLY NUMBER OF EARNERS ESTIMATED FROM MONTHLY EMPLOYMENT

This analysis of the number of earners reported in population-based income surveys as well as the number of persons reported as having done any work for wages or salary or on their own

<sup>2</sup> The monthly labor force enumeration in the Current Population Survey is for the calendar week including the eighth of each month. In view of the short-run, sporadic, and isolated instances of employment which give rise to most of the underreporting, it seems reasonable to assume that the rate of underreporting would be raised somewhat if all weeks could be checked.

account during a year clearly indicates the necessity of developing independent estimates of the number of earners.

The number of earners can be estimated from annual estimates of employment based on monthly enumerations. Because households remain in the Census Bureau Current Population Survey sample for several months histories of employment status can be constructed for periods ranging from 6 to 9 months.

The Bureau of the Census has recently completed a special tabulation for a sample enumerated from December 1946 through August 1947. Adding to the number employed in the first month the number who first became employed in each succeeding month yielded gross employment for the nine months. Projecting the increments to gross employment for September, October, and November 1947 yields gross employment for the twelve months, December 1946 to November 1947: 68 million in the civilian noninstitutional population 14 and over in November 1947 had been employed and received earnings in at least one of the twelve survey weeks. The figures for the calendar year 1947 would be slightly higher since the trend of employment was upward—probably approximately equal to the 69 million estimate outlined earlier. Like the earlier estimate it has not been adjusted for the additional earners that would be reported if all weeks were surveyed.

#### D NECESSITY FOR PREPARING A SEPARATE INCOME SIZE DISTRIBUTION FOR EARNERS NOT REPORTING EARNINGS IN SURVEYS

The impossibility of obtaining good earnings size distributions merely by applying the percentage distribution of those reporting earnings to an independent estimate of the number of earners seems clear. The earnings size distribution of those who do not report earnings is quite different from the distribution of those who report earnings. Experience with questions on annual employment asked in connection with the Current Population Survey indicated that most of those who failed to report employment had done little work. Obviously casual employment or one or two brief spells of employment during a year would be likely to be forgotten. It seems probable also that some re-

spondents say they have no earnings when they are unwilling to report. Some of those appearing as zero earnings respondents are, therefore, concealed refusals; a significant proportion are likely to be in the higher income brackets.

The information available does not provide a basis for estimating the size distribution of earnings for those failing to report earnings. Two approaches suggest themselves: (a) to develop such data by matching Census reports with OASI and BIR reports; (b) to conduct intensive follow-up interviews with a sample of persons for whom no earnings are reported in an income survey. Both approaches could be used, of course, in measuring the extent of income underreporting as well as the earnings or income of those reporting no earnings or income. Either project would be expensive and both would present operating and public relations problems. Furthermore, neither could yield very satisfactory results. Matching field survey schedules with OASI records would not provide a check on earnings in domestic service, agriculture, and other work not covered. Much of the income that is not reported in field surveys would not be reported for income tax purposes because of its small amount, its illegal source, or for other reasons. Experience indicates that intensive follow-up interviews would likely unearth only part of the income not reported in the first interview, and would be primarily a measure of variability of response rather than a factual basis for adjusting the income distributions obtained in the original interview.

Notwithstanding the shortcomings and difficulties of these methods, it would seem desirable to push forward exploratory work along these lines, as well as to seek to develop other methods. Pending the result of such work, it will be necessary to make rather arbitrary assumptions how unreported earnings are distributed by size.

Data on income distributions obtained in field surveys provide much valuable information, notwithstanding the extent of underreporting of both the number of income recipients and the amount received. Used in connection with statistics from other sources, they are an indispensable part of the information essential for deriving series on size distributions of income.

