

Susan Palmer, The Bureau of the Census

Introduction

An often frustrating component of all sample surveys which rely on public cooperation is the element of nonresponse. ^{1/} Failure to obtain observations from all members of the sample can lead to appreciable biases in the estimates and faulty inferences about the population. The character and influence of the nonresponse component of sample surveys are as varied as the surveys themselves. Their nature often depends on the subject of inquiry, methods of data collection, interviewer ability and motivation, public opinion, allotted time for survey completion, response rates, and even on the changing seasons. Survey methodologists have adopted a variety of methods to compensate for nonresponses. Some of these methods are rather sophisticated, such as setting up response models which take into account the probability that a household will be contacted and interviewed, or subsampling the nonresponse cases and carrying out intensive follow-up on the subsample to collect data for the nonresponse group. Other methods are more straightforward. The adjustment procedure used most often is an inflation of the interview universe, based on the assumption that the unknown, or nonresponse, elements are like the known, or response, elements.

Whether nonresponse presents a serious limitation on the total estimate is principally related to the nonresponse group and to the accuracy desired in the survey results. For surveys with moderate response rates where only crude measures are desired, the potential biases of nonresponse may be of little importance if what is known about the nonresponse elements indicates that they do not differ markedly from the response elements. Where a high degree of precision is desired, however, even a relatively low nonresponse rate can seriously affect the survey results.

The present paper concerns research conducted in connection with the Current Population Survey, a monthly household survey which is the source of the official estimates of employment and unemployment. As part of the effort to improve the accuracy of the CPS estimates, the Census Bureau has been actively concerned with the nonresponse problem - its size, its effect on the statistics,

and how to improve the adjustment procedures for nonresponse. Realizing that serious biases can arise from differences between interviewed and nonresponse households, the Census Bureau has always placed great emphasis on keeping the CPS nonresponse rate at a minimum consistent with budget and time considerations (interviewing must be completed in one week). As a result, the nonresponse rate in the Current Population Survey averages around $4\frac{1}{2}$ percent, and ranges from a low of 3 to $3\frac{1}{2}$ percent in certain spring and fall months to a high of about 6 percent in some summer months--one of the lowest nonresponse rates among surveys where public cooperation is elicited.

As part of a broader study begun in 1963, observations on nonresponse households were obtained by carrying out an intensive field follow-up of such households during the three weeks following the CPS survey week in September 1965. Unfortunately, interviews were obtained for only about one-half of the survey households in the scope of this study. Because such a large proportion of these households remained inaccessible, the data presented here are themselves biased estimates for the nonresponse group. However, the results are consistent with what one would logically expect of CPS nonresponse, and probably understate the true differences between the nonresponse and response households. That is, the "hard core" nonresponse elements probably differ even more markedly, and in the same direction, than those for which observations were obtained.

The Character of Nonresponse Households

The information gathered on the CPS nonresponse households provides some insight into the character of these households and the persons within. Obviously, the probability of finding a qualified respondent at home at any time is related to the size of the household. It is not surprising that more than 1 in 4 of the nonresponse units are single-person households, as compared with a total household figure of only 1 in 7. The average number of persons 14 years and over in the response household in CPS is 2.3, while the nonresponse units average only about 2 such persons (See Table 1). The difference between the two groups is even more striking for total household size--3.4 persons per household for the interviewed units versus 2.6 persons for the nonresponse groups. We are not directly concerned in labor force statistics with children under 14 years of age. However, data on family characteristics, health, etc., from supplemental CPS questions do take into account the younger age groups.

^{1/} We are not concerned in this paper with the problems of incomplete information (e.g., where a sample unit is interviewed but some of the desired information is omitted, either inadvertently or for other reasons) but with only the problems of nonresponse caused by the refusal of the sample unit to cooperate or by its unavailability to the survey interviewer.

Nonresponse persons show a slightly higher median age than the interviewed groups (43 years as compared to about 40 years), caused by a conspicuous shortage of persons in the younger age categories. Almost 17 percent of the interviewed persons in CPS are between the ages of 14 and 19 years while only 12 percent of the nonresponse persons are of these ages (see Table 2).

Household sizes and population age distributions are relatively static, that is, households will exhibit the same general size and age distributions whether they are response or non-response in any month. The important point is that certain households, because of their composition, are more likely to be nonresponse than others.

With these demographic differences it is reassuring to find the labor force participation rate for the response and nonresponse groups about the same, and along with it the overall employment and unemployment rates (see Table 3). Large differences exist, however, in two labor force categories; persons working part time and those with a job, but not at work. About 17 percent of the interviewed labor force are part-time workers, whereas the comparable rate for nonresponse persons is only 12 percent. But even more striking is that while 4 percent of the interviewed labor force participants are temporarily away from their work for the survey period, about 10 percent of the nonresponse persons fall into this category. It is unlikely that these labor force differences can be attributed to the demographic composition of the non-response households. By comparing the labor force status of identical persons in response and nonresponse months we discovered that a definite change in the working routine had frequently taken place. 2/ Over half of the persons "with a job, not at work" during the nonresponse month had been working full time during a previous response month. Clearly, a large proportion of the nonresponse universe had shifted from "working" into the rather miscellaneous category composed of vacationers, leave-takers, and the ill.

A similar change is associated with nonresponse persons outside of the labor force. While 22 percent of the non-labor force interviewed persons are attending school during survey week, only 16 percent of the nonresponse persons are

thus engaged at the time of the survey. Partially explained by the shortage of school-age children in the nonresponse group, the difference is also attributable to an influx of school-attending persons into another heterogeneous category, "other." The "other" category appears to function in much the same way for nonresponse persons outside the labor force as the "with a job, not at work" category does for the nonresponse labor force participants. This non-labor force classification includes students temporarily away from class, housewives whose week's activities were not as they normally are, etc. To illustrate, consider the categorical changes in both labor force and non-labor force status of a family during a vacation month. We conclude, therefore, that certain selective shifts in status are associated with CPS households being nonresponse. They are not like the response households, nor are they like themselves, that is, they do not have the same survey characteristics that they had in the months in which they were interviewed. (This finding has a significant effect on proposed nonresponse adjustment procedures to be discussed later).

THE UNAVAILABLES. The CPS nonresponse elements are of two basic types--those households which were not interviewed because they were unavailable during the survey time period, and those which, although contacted, would not cooperate with the survey. Mainly for the purpose of interviewing control the unavailable nonresponses are customarily divided into three groups on the basis of their reason for nonresponse. These are the "no one home," "temporarily absent," and "other unavailable" households. Since collectively these households normally account for about two-thirds of the total CPS nonresponse component, their individual natures, to a large extent, are responsible for the differences previously mentioned in the known and unknown universes.

The No One Home Households: This group represents those households which cannot be found at home by the interviewer after repeated calls at varying times. Probably related to their unavailability, the no one home persons 14 years and over have the highest labor force participation rate of any nonresponse group--about 68 percent of this population. 36 percent of the no one home households are single person, and the average number of persons 14 and over in these units is only 1.8.

The no one home rate of nonresponse fluctuates between 1 and 1½ percent of the total CPS households throughout the year.

2/ The Current Population Survey is a rotation panel operation. Households are in sample for four consecutive months, drop out for the next eight months, and then return for four more months. 75 percent of the households in sample in any one month were in sample the preceding month.

The Temporarily Absent Households: The households of this type, unlike the no-one-homes, have no probability of being interviewed. The interviewer knows that they will be unavailable for the entire survey period. The temporarily absent households contain only 1.7 persons 14 years and over. In contrast to the no-one-homes, the persons in these households have an uncommonly low labor force participation rate (only 14 percent), and an unusually high percentage of them are retired (reflected in the non-labor force category of "other"). Most important, however, is the large proportion of persons "with a job, not at work"--about 28 percent of all temporarily absent persons in the labor force. These households appear to create the large difference between response and non-response with respect to this category. Their influence on the total component is greatest during the summer months when they account for almost $\frac{1}{2}$ of all nonresponse cases.

The Other Unavailable Households: Representing only $\frac{1}{2}$ of 1 percent of the total CPS sample, the other unavailable households have a negligible effect on the character of the nonresponse component. The most heterogeneous of the nonresponse types, this group consists of households which normally would have been interviewed except for extenuating circumstances (impassable roads, a death in the family, etc.). The only noteworthy attribute of its diverse members is their abnormally high "with a job, not at work" rate. Thus they contribute, however slightly, to the character of the total nonresponse universe.

THE UNCOOPERATIVES. Information on the characteristics of uncooperative nonresponse (about $1\frac{1}{2}$ percent of the total CPS sample) was not collected in the manner used for the "unavailables." No attempt was made through follow-up personal interview to persuade the refusing persons to cooperate. Instead it seemed reasonable to assume no inherent labor force change associated with their refusing to be interviewed, but to regard their uncooperativeness as an individual factor--a reflection of the respondent's attitudes and personal feelings, not of his labor force status during the month in question. Therefore, demographic and labor force information for this subgroup of the population was obtained from months in which refusal households did cooperate and used as an indication of refusal household character. ^{3/}

^{3/} Not reflected in the refusal data are those households which were never interviewed during their 8 months in sample. This necessary omission of information represents only about one-tenth of one percent of the total CPS sample.

The Refusal Households: We have yet to uncover any definitely correlated nonresponse characteristics for the refusal group. Whereas an obvious difference in household size was noted for the unavailable nonresponse types, in particular the no-one-homes and temporarily absents, the refusal household size of 2.3 persons 14 years and over is the same as the interview statistic. Also, the labor force status of the uncooperative persons in months when they were interviewed closely resembles the status of persons who had never refused during their eight months in sample.

The level of refusal nonresponse shows very little seasonal variation. Its effect on the total nonresponse component in CPS is static, and because of its close resemblance to the interviewed universe, actually serves to temper the effect of the unavailables and bring the total character of nonresponse closer to that of the interviewed.

The Influence of Nonresponse on Labor Force Statistics

When the survey results are in, and estimates are being made for the universe, the nonresponse cases must be taken into account either implicitly or explicitly. If the nonresponses are apparently disregarded by inflating the interview data to the sample frame, they have, in effect, been treated as though they were a representative sample of the universe. Thus, even the lack of an explicit method of dealing with nonresponse can be considered an adjustment procedure.

The nonresponse adjustment technique used in the Current Population Survey is a differential weighting system applied to the interview data by color and residence within each of about 70 groups of sample units on the assumption that the average population size and the labor force status of the interview and nonresponse households are identical within these groups. ^{4/} Although this assumption does not introduce appreciable biases in the major labor force categories such as employed and unemployed, the substitution of actual nonresponse characteristics for those "manufactured" to represent the nonresponse group brings to light troublesome differences in some of the minor categories (see Table 4). For example the difference between the published and the "true" proportion of persons "with a job,

^{4/} For a detailed explanation of the CPS non-response adjustment procedure, see Part VII of Technical Paper No. 7--"A Report on Methodology", Bureau of the Census, Dept. of Commerce, 1963.

not at work" is two-tenths of one percent, or about 5 percent of the published estimate. As the approximate standard error on this rate for the interview cases is about one-tenth of one percent, the imputation of interview characteristics for nonresponse households has substantially affected the reliability of the "with a job, not at work" category.

Nonresponse biases of about the same level as the standard error are noted for part time workers, school attenders, and "other--not in the labor force." Clearly, the unusually low level of nonresponse in CPS does not preclude the introduction of errors by the assumption of response and nonresponse comparability. One way of viewing this effect is to note that if the nonresponse bias could be eliminated, the sample size could be cut in half without increasing the total mean square error for these items (although obviously a decision concerning sample size in CPS would not be based only on consideration of the specific items that are seriously affected by nonresponse).

Several different imputation techniques for nonresponse were explored to determine whether any others would decrease this nonresponse bias in the sample estimates. First, we considered the substitution of labor force data from a previous interview month for current nonresponse households. The assumption here is that a household tends to exhibit similar labor force characteristics over a series of months independent of its interview status. As discussed previously,

this assumption for nonresponse households is in error, and labor force estimates derived in this manner do not seem to represent any improvement over those derived by the present adjustment method.

Secondly, it seemed intuitively reasonable to expect the nonresponse units in CPS to resemble more closely the "hard to enumerate" households than the interview universe as a whole, the notion being that the substitution of a portion of the response data for the nonresponse might be more accurate than the present method of inflation. However, when comparisons were made between the labor force status of the households not contacted until the second or third personal visit and those never enumerated, the important differences in the nonresponse-related categories of "with a job, not at work," "school," etc., still existed.

Different statistics are likely to be affected in vastly different ways by nonresponse. As seen in the results of the CPS research, the fact that the estimates of primary importance are not seriously biased did not assure that all labor force statistics would not be affected. Nonresponse households do not constitute a single, homogeneous group. It is therefore difficult to anticipate the problems that they will cause in a particular survey. We do not presume that the findings presented here on the phenomenon of nonresponse apply to the general nonresponse universe, but are offered in an effort to provide some insight into the relationship between the known and unknown in sample surveys.

Table 1: Size of Household for the Total Population of the United States and for Nonresponse Households in the Current Population Survey, September 1965.

Size of household	Total house-holds <u>a/</u>	Total non-response households <u>b/</u>	Unavailable nonresponse <u>c/</u>			Uncooperative nonresponse <u>d/</u> Refusals
			No one home	Tempo-rarily absent	Other unavail-able	
All households	100.0	100.0	100.0	100.0	100.0	100.0
1 person	13.1	28.5	36.8	34.1	31.2	16.4
2 persons	27.8	35.0	33.3	42.2	31.2	32.4
3 persons	18.9	13.8	10.2	11.6	10.1	19.6
4 persons	17.6	10.9	8.4	7.0	16.5	14.2
5 persons	11.5	5.3	4.2	2.7	4.6	8.3
6 persons or more	11.1	6.5	7.0	2.3	6.4	9.2
Average number of persons per household	3.4	2.6	2.3	2.1	2.5	3.0
Average number of persons 14 years of age and older per household	2.3 <u>e/</u>	2.0	1.8	1.7	2.0	2.3

a/ Source: Current Population Reports, Series P-20, No. 106, Table 3; March 1960.

b/ Estimates obtained by weighting the unavailable and the uncooperative households to the nonresponse sample level for September 1965.

c/ Data obtained from CPS records of households which were unavailable during the month of September 1965 but had been interviewed during a previous month. Unweighted sample totals for No one home - 285; Temporarily absent - 258; Other unavailable - 109.

d/ Data obtained from CPS records of households in sample as of December 1965 which, during their 8 months in sample, refused to cooperate in at least one month, and were interviewed in at least one month. Unweighted sample total - 720 households.

e/ Estimate the same in both March 1960 and September 1965 for total CPS households.

Table 2: Age Distribution for Persons 14 Years and Older in Response and Nonresponse Households in the Current Population Survey, September 1965.

Age	Total house- holds <u>a/</u>	Total non- response households <u>b/</u>	Unavailable nonresponse <u>c/</u>			Uncooperative nonresponse <u>d/</u> Refusals
			No one home	Tempo- rarily absent	Other unavail- able	
Persons 14 years of age and older	100.0	100.0	100.0	100.0	100.0	100.0
14-19 years	16.7	12.4	13.9	6.7	12.1	14.8
20-24 years	8.8	9.9	15.1	8.2	7.5	8.0
25-29 years	8.0	8.3	12.5	6.9	6.5	6.8
30-34 years	7.9	6.5	7.6	4.8	9.5	6.0
35-39 years	9.1	7.0	7.0	5.6	3.5	8.6
40-44 years	9.5	9.0	8.3	5.9	10.6	10.7
45-49 years	8.4	8.8	9.1	5.0	11.1	10.2
50-54 years	7.7	8.8	7.4	7.8	11.1	9.8
55-59 years	6.7	8.6	7.8	10.2	8.5	8.3
60-64 years	5.4	7.0	5.0	10.6	5.5	6.8
65-69 years	4.8	5.7	2.8	10.2	8.0	4.6
70+ years	8.1	7.8	3.6	18.0	6.0	5.4
Median age	39.7	43.2	35.6	54.4	45.1	42.7

a/ Estimates include only those persons in households which were interviewed in September 1965. The CPS adjustment for nonresponse households is not reflected in the estimates.

b/ Estimates obtained by weighting the unavailable and the uncooperative households to the nonresponse sample level for September 1965.

c/ Data obtained from CPS records of households which were unavailable during the month of September 1965 but had been interviewed during a previous month. Unweighted sample totals for No one home - 503; Temporarily absent - 461; Other unavailable - 199.

d/ Data obtained from CPS records of households in sample as of December 1965 which, during their 8 months in sample, refused to cooperate in at least one month, and were interviewed in at least one month. Unweighted sample total - 1,658 persons.

Table 3: Labor Force Characteristics for Persons 14 Years of Age and Older in Response and Nonresponse Households in the Current Population Survey, September 1965

Labor Force Status	Response persons <u>a/</u>	Total non-response persons <u>b/</u>	Unavailable nonresponse <u>c/</u>			Uncooperative nonresponse <u>d/</u> Refusals
			No one home	Temporarily absent	Other unavailable	
Persons 14 years of age and older	100.0	100.0	100.0	100.0	100.0	100.0
In labor force	56.3	56.6	67.7	44.0	51.7	57.4
Not in labor force	43.7	43.4	32.3	56.0	48.3	42.6
<u>In labor force</u>	100.0	100.0	100.0	100.0	100.0	100.0
Employed	96.0	96.7	96.2	94.7	98.6	97.3
Working 35+ hours	75.0	74.2	78.8	63.2	68.9	76.7
Working 1-34 hours	17.0	12.1	12.3	3.5	14.9	14.9
With a job, not at work	3.9	10.3	5.2	28.1	14.9	5.7
Unemployed	4.0	3.3	3.8	5.3	1.4	2.7
<u>Not in labor force</u>	100.0	100.0	100.0	100.0	100.0	100.0
Keeping house	59.6	57.8	54.5	62.8	55.1	56.9
School	22.4	16.1	16.8	5.5	24.6	20.5
Other	15.2	22.4	22.8	29.7	15.9	19.3
Unable to work	2.9	3.6	5.9	2.1	4.3	3.3

- a/ Estimates include only those persons in households which were interviewed in September 1965. The CPS adjustment for nonresponse households is not reflected in the estimates.
- b/ Estimates obtained by weighting the unavailable and the uncooperative households to the nonresponse sample level for September 1965.
- c/ Estimates are based on persons in households which were unavailable to be interviewed during survey week in September 1965 but were interviewed during a three week follow-up period (approximately 50 percent of the unavailable nonresponse households were interviewed during the follow-up). Unweighted sample totals for No one home - 313; Temporarily absent - 259; Other unavailable - 143.
- d/ Data obtained for households which refused to cooperate at least once during the months of June through November 1964 and were interviewed at least once either before or after the refusal month. Unweighted sample total - 5,967 persons.

Table 4: Labor Force Characteristics for Persons 14 Years of Age and Older Using the Present CPS Nonresponse Adjustment Method, and as Estimated Substituting Actual Data for the Nonresponse Households in the Current Population Survey, September 1965.

Labor Force Status	Estimated using present nonresponse adjustment method <u>a/</u>	Substituting actual data for nonresponse households <u>b/</u>	Absolute difference between Columns (1) and (2)	Approximate standard error on level of estimated percent in Column (1)
	(1)	(2)	(3)	(4)
Persons 14 years of age and older	100.00	100.00		
In labor force	56.24	56.27	0.03	
Not in labor force	43.76	43.73	0.03	
<u>In labor force</u>	100.00	100.00		
Employed	95.97	96.00	0.03	0.1
Working 35+ hours	75.04	75.00	0.04	0.2
Working 1-34 hours	17.02	16.88	0.14	0.2
With a job, not at work	3.92	4.12	0.20	0.1
Unemployed	4.03	4.00	0.03	0.1
<u>Not in labor force</u>	100.00	100.00		
Keeping house	59.57	59.50	0.07	0.2
School	22.38	22.18	0.20	0.2
Other	15.15	15.40	0.25	0.2
Unable to work	2.89	2.92	0.03	0.1

a/ The present nonresponse adjustment method is based on the assumption that nonresponse elements within a specific region-color-residence category are like the response elements in that category, and the response elements are weighted accordingly. For a detailed explanation see Part VII of Technical Paper No. 7, "The Current Population Survey--A Report on Methodology," Bureau of the Census, U.S. Dept. of Commerce, 1963.

b/ Data were collected for nonresponse households in a special study in September 1965. These data were used as an estimate for the nonresponse households and combined with data from response households to obtain this estimate.