

THE QUALITY OF REPORTING SOCIAL SECURITY NUMBERS IN TWO SURVEYS

Mitsuo Ono, George F. Patterson, and Murray S. Weitzman*
U.S. Bureau of the Census

Introduction

This analysis presents findings on the rate of reporting and the extent of "correct" reporting of Social Security numbers (hereafter designated as SSN's in this paper) by persons included in a sample of households obtained in two mail-out/mail-back content pretest surveys. These pretest surveys were conducted by the Bureau of the Census in May 1966 and April-May 1967 in connection with the planning of the 1970 Census of Population and Housing. In May 1966, the Bureau of the Census conducted the First Content Pretest in two areas--St. Louis Park, Minnesota and Yonkers, New York. This pretest survey universe contained approximately 4,700 households. A simple random sample of approximately 660 households (or about 14 percent of the total number of households) was selected for this study on SSN's. These households included about 2,000 persons of all ages. In April and May 1967, the Bureau of the Census conducted another pretest survey (the Second Content Pretest) at Gretna, Louisiana. The pretest survey universe was approximately 4,200 households. A simple random sample of 765 households was selected for this analysis. These households contained about 2,500 persons of all ages.

In both of these pretest surveys, all persons in the surveyed households who had an SSN (or a Railroad Retirement number) were requested to write in their SSN in the pretest schedules. Whenever a member of the household had no SSN, he was requested to fill in the "none" circle located directly to the right of the space for the nine-digit SSN. Respondents were requested not to guess their numbers, but to consult records, such as Social Security account cards or Federal tax records, whenever necessary. Thus, the col-

lection of SSN's presented somewhat different reporting problems as compared with other Census questionnaire items that could be estimated without the need to consult records.

In this study, data intended for statistical tabulations were transcribed from the schedules onto special forms, verified, and forwarded to the Social Security Administration for processing. Statistical tabulations on the correctness of the SSN information were obtained from the Social Security Administration. Selected data from these statistical tabulations are presented in the appendix tables.

Summary of Findings

Among all persons 14 years of age or over, approximately 76 percent reported an SSN in the two pretest surveys. Fourteen percent marked the "none" circle indicating that they did not have an SSN and about 10 percent did not answer the SSN question. Approximately 77 percent of all persons 14 years old or over answered the question "correctly." This "correct" rate consisted of 67 percent of cases reporting that they had an SSN and 10 percent of cases reporting that they did not have an SSN. In other words, among persons reporting that they had an SSN, i.e., 76 percent, approximately 87 percent reported their "correct" SSN. For persons who reported they had no SSN's, i.e., 14 percent, 73 percent reported "correctly," that is, no SSN's were located for such persons.

For some of the persons reporting an incorrect SSN, no SSN, or who did not answer the SSN question, it was possible to locate their "correct" SSN's through additional effort. This effort resulted in improving the rate of "correct" response from 77 percent to 88 percent.

*Comments represent views of the authors and not necessarily those of the Bureau of the Census.

Note: Findings from another statistical study on Social Security numbers can be found on pages 18 and 19 in the following unpublished paper: "Interacting Data Systems and the Measurement of

Income Size Distributions," by Joseph Steinberg, Social Security Administration, presented at the Conference on Research in Income and Wealth, Philadelphia, Pennsylvania, March 24-25, 1967. Suggestions by Mr. Joseph Steinberg of the Social Security Administration are gratefully acknowledged.

Table A.--REPORTING OF SSN's BY PERSONS 14 YEARS OLD AND OVER IN THE FIRST AND SECOND CONTENT PRETEST SURVEYS

Response	Total	Answered question			"Incorrect" answer possibly "corrected" by search	Residual (col. 3 minus col. 5)
		"Correctly"	"Incorrectly"	Information not available		
	(1)	(2)	(3)	(4)	(5)	(6)
All.....	2,962 (100.0)	2,276 (76.8)	676 (22.8)	10 (0.3)	332 (11.2)	344 (11.6)
Reported having SSN.....	2,257 (76.2)	1,974 (66.6)	283 (9.6)	--	88 (3.0)	195 (6.6)
Reported <u>not</u> having SSN.....	413 (13.9)	302 (10.2)	101 (3.4)	10 (0.3)	101 (3.4)	--
Did not answer.....	292 (9.9)	--	292 (9.9)	--	143 (4.8)	149 (5.0)

Note: Percentages in parentheses.

Source: Appendix tables 1 and 2.

Frequency of Reporting SSN's

The usefulness of obtaining SSN information in survey work depends on the percent of the population reporting SSN's. As shown in table B, persons 25 to 64 years of age reported that they had SSN's more frequently than did persons in the under 25-year or over 65-year age groups. A comparison of reporting rates of persons 14 to

19 years of age with that of persons 20 years old and over included in the Second Content Pretest survey is presented in table C. About 40 percent of persons 14 to 19 years of age reported not having an SSN as compared with 12 percent of persons 20 years old and over.

Table B.--REPORTING OF SSN's BY PERSONS 14 YEARS OLD AND OVER BY AGE GROUPINGS IN THE FIRST AND SECOND CONTENT PRETEST SURVEYS

Response	Total	Age groups					
		14 - 24	25 - 44	45 - 54	55 - 64	65 - 71	72 and over
Total.....	2,962 (100.0)	789 (100.0)	921 (100.0)	493 (100.0)	398 (100.0)	192 (100.0)	169 (100.0)
Reported having SSN.....	2,257 (76.2)	492 (62.4)	774 (84.0)	402 (81.5)	323 (81.1)	144 (75.0)	122 (72.2)
Reported <u>not</u> having SSN.....	413 (13.9)	209 (26.5)	56 (6.1)	34 (6.9)	48 (12.1)	36 (18.7)	30 (17.8)
Did not answer.....	292 (9.9)	88 (11.2)	91 (9.9)	57 (11.6)	27 (6.8)	12 (6.2)	17 (10.1)

Source: Appendix table 2.

Table C.--REPORTING OF SSN's BY PERSONS 14 TO 19 YEARS OF AGE AND PERSONS 20 YEARS OLD AND OVER IN THE SECOND CONTENT PRETEST SURVEY ONLY

Response	Percent	
	14 to 19 years	20 years old and over
Total.....	100.0	100.0
Reported having SSN.....	46.4	76.7
Reported <u>not</u> having SSN.....	39.6	12.2
Did not answer.....	14.0	11.0

Source: Special tabulations not included in appendix tables.

Table D shows that male persons 14 years old and over report having SSN's more frequently than female persons.

Table D.--REPORTING OF SSN's BY MALE AND FEMALE PERSONS, 14 YEARS OLD AND OVER, IN THE FIRST AND SECOND CONTENT PRETEST SURVEYS

Response	Total	Male	Female
Total.....	2,962 (100.0)	1,386 (100.0)	1,576 (100.0)
Reported having SSN.....	2,257 (76.2)	1,150 (83.0)	1,107 (70.2)
Reported <u>not</u> having SSN.....	413 (13.9)	96 (6.9)	317 (20.1)
Did not answer.....	292 (9.9)	140 (10.1)	152 (9.6)

Source: Appendix table 3.

Information in table E obtained from the Second Content Pretest survey indicates that household heads report more frequently having SSN's than "other" household members. About 86

percent of household heads reported having SSN's as compared with 62 percent for other members of households.

Table E.--REPORTING OF SSN's BY HEADS OF HOUSEHOLDS AND OTHER MEMBERS OF HOUSEHOLDS, 14 YEARS OLD AND OVER, IN THE SECOND CONTENT PRETEST SURVEY ONLY

Response	Total	Heads of households	Other members of households
Total.....	1,655 (100.0)	723 (100.0)	932 (100.0)
Reported having SSN.....	1,194 (72.1)	620 (85.8)	574 (61.6)
Reported <u>not</u> having SSN.....	271 (16.4)	36 (5.0)	235 (25.2)
Did not answer.....	190 (11.5)	67 (9.3)	123 (13.2)

Source: Appendix table 4.

"Correct" Reporting of SSN's

Persons reporting "correctly" their SSN's comprise two subgroups: (1) persons not only reporting having SSN's but also reporting them "correctly" and (2) persons reporting "correctly" not having SSN's. The proportion of "correct" reporting can be computed using either a population base including all persons 14 years of age and over in the pretest surveys or only persons 14 years of age and over in the particular population group to be analyzed. The proportion of "correct" reporting using the base covering all

persons 14 years and over is found in table A. The proportions of "correct" reporting using the base covering only persons in the specified group are presented in tables F, G, and H, below.

Table F shows the proportion of "correct" reporting of persons reporting having SSN's and persons reporting not having SSN's. The proportions of persons who reported "correctly" were about the same for the various characteristics shown below except for persons in the age category 72 years and older.

Table F.--CORRECT REPORTING OF SSN's BY PERSONS 14 YEARS OLD AND OVER WHO REPORTED HAVING AND NOT HAVING SSN's IN THE FIRST AND SECOND CONTENT PRETEST SURVEYS

Selected characteristics	Reported having SSN and not having SSN	"Correctly" reported	Percent "correct" reporting
	(1)	(2)	(3) (2)/(1) X 100
All ^{1/}	2,670	2,276	85.2
<u>By age</u> ^{1/}			
14 to 24 years.....	701	583	83.2
25 to 44 years.....	830	720	86.7
45 to 54 years.....	436	385	88.3
55 to 64 years.....	371	316	85.2
65 to 71 years.....	180	155	86.1
72 years and over.....	152	117	77.0
<u>By sex</u> ^{1/}			
Male.....	1,246	1,074	86.2
Female.....	1,424	1,202	84.4
<u>By relationship</u> ^{2/}			
Heads of households.....	656	562	85.7
Other members of households.....	809	682	84.3

^{1/} Results obtained from both pretest surveys.

^{2/} Results obtained from the Second Content Pretest survey only.

Source: Appendix tables 2, 3, and 4.

Table G demonstrates that persons willing to report SSN's tend to report their SSN's "correctly." The proportion of persons reporting

correctly did not vary much by different age groups, except for persons in the age group 72 years old and over.

Table G.--REPORTING OF "CORRECT" SSN's BY PERSONS 14 YEARS OLD AND OVER WHO REPORTED SSN's IN THE FIRST AND SECOND CONTENT PRETEST SURVEYS

Selected characteristics	Reporting SSN	Reporting a "correct" SSN	Percent "correct" reporting
	(1)	(2)	(3) (2)/(1) X 100
All ^{1/}	2,257	1,974	87.5
By age ^{1/}			
14 to 24 years.....	492	436	88.6
25 to 44 years.....	774	679	87.7
45 to 54 years.....	402	358	89.1
55 to 64 years.....	323	280	86.7
65 to 71 years.....	144	124	86.1
72 years and over.....	122	97	79.5
By sex ^{1/}			
Male.....	1,150	1,010	87.8
Female.....	1,107	964	87.1
By relationship ^{2/}			
Heads of households.....	620	537	86.6
Other members of households.....	574	511	89.0

1/ Results obtained from both pretest surveys.

2/ Results obtained from the Second Content Pretest survey only.

Source: Appendix tables 2, 3, and 4.

Table H shows that among persons reporting not having SSN's, about 73 percent reported the question "correctly." Female persons and "other"

members of the household who reported not having SSN's tend to report this item more correctly than male persons and household heads.

Table H.--REPORTING OF "CORRECT" SSN's BY PERSONS 14 YEARS OLD AND OVER WHO REPORTED NOT HAVING SSN's IN THE FIRST AND SECOND CONTENT PRETEST SURVEYS

Selected characteristics	Reported <u>not</u> having SSN	No SSN was found	Percent "correct" reporting
	(1)	(2)	(3) (2)/(1) X 100
All ^{1/}	413	302	73.1
By age ^{1/}			
14 to 24 years.....	209	147	70.3
25 to 44 years.....	56	41	73.2
45 to 54 years.....	34	27	79.4
55 to 64 years.....	48	36	75.0
65 to 71 years.....	36	31	86.1
72 years and over.....	30	20	66.7
By sex ^{1/}			
Male.....	96	64	66.7
Female.....	317	238	75.1
By relationship ^{2/}			
Heads of households.....	36	25	69.4
Other members of households.....	235	171	72.8

1/ Results obtained from both surveys.

2/ Results obtained from the Second Content Pretest survey only.

Source: Appendix tables 2, 3, and 4.

Information Obtained From Additional Effort

Data in table I show the extent of additional SSN's located for respondents who either reported an "incorrect" SSN, or reported no SSN, or who did not answer the question. Among 283 persons who had reported "incorrect" SSN's, possible "correct" SSN's were found for about a third of these persons. Among 413 persons who had reported that they had no SSN's, possible correct SSN's were designated for about 25 percent of these persons. Among 292 persons who did not answer the question, possible SSN's were found for about half of these persons. Approx-

mately two-thirds of the number of persons in the age category 14 to 24 years who did not answer the question were found to have possible correct SSN's. Among persons either reporting "incorrectly" or not answering the question, possible SSN's were located more frequently among male persons than for female persons. However, possible SSN's were located at about the same rate for both household heads and other members of households who either answered this question improperly or omitted the item.

Table I.--POSSIBLE SSN'S LOCATED FOR PERSONS REPORTING SSN "INCORRECTLY," NOT HAVING AN SSN, OR NOT ANSWERING SSN QUESTION IN THE FIRST AND SECOND CONTENT PRETEST SURVEYS

Response	Total ^{1/}	By age						By sex		By relationship ^{2/}	
		14 to 24	25 to 44	45 to 54	55 to 64	65 to 71	72 and over	Male	Female	Heads of households	Other members of households
A. Persons reporting "incorrect" SSN.....	283	56	95	44	43	20	25	140	143	83	63
B. "Incorrect" answers possibly corrected.....	88	17	30	16	16	4	5	52	36	25	18
C. (B/A) net percentage.....	31.1	30.4	31.6	36.4	37.2	20.0	20.0	37.1	25.2	30.1	28.6
D. Persons reporting not having SSN.....	413	209	56	34	48	36	30	96	317	36	235
E. Possible SSN located.....	101	56	12	7	12	5	9	29	72	11	64
F. (E/D) net percentage.....	24.5	26.8	21.4	20.6	25.0	13.9	30.0	30.2	22.7	30.6	27.2
G. Persons not answering questions.....	292	88	91	57	27	12	17	140	152	67	123
H. Possible SSN located.....	143	58	44	22	9	4	6	75	68	33	57
I. (H/G) net percentage.....	49.0	65.9	48.4	38.6	33.3	33.3	35.3	53.6	44.7	49.3	46.3
J. All persons 14 years and over..	2,962	789	921	493	398	192	169	1,386	1,576	723	932
K. Possible SSN located from search (B + E + H).....	332	131	86	45	37	13	20	156	176	69	139
L. (K/J) net percentage.....	11.2	16.6	9.3	9.1	9.3	6.8	11.8	11.3	11.2	9.5	14.9
M. "Correct" reporting rate without search ^{3/}	76.8	73.9	78.2	78.1	79.4	80.7	69.2	77.5	76.3	77.7	73.2
N. "Correct" reporting rate with search (L + M).....	88.0	90.5	87.5	87.2	88.7	87.5	81.0	88.8	87.5	87.2	88.1

^{1/} Results from First and Second Content Pretest surveys.

^{2/} Results from Second Content Pretest survey only.

^{3/} Lines C through J, col. 3 + col. 9 (table 2).

Line C + line G (table 3).

Line C + line G (table 4).

Overall, possible "correct" SSN's were located for 332 persons, or for approximately 11 percent of the 2,962 persons 14 years of age and over included in both pretest surveys. This additional effort resulted in increasing the accuracy of the cases sampled from 77 to 88 percent. The highest gain in the improvement in the accuracy of reporting occurred in the 14 to 24-year-old age group.

Summary

In general, it can be concluded from this analysis that persons who are willing to report SSN's tend to report them correctly. The "best quality" of reporting SSN's can be expected for male heads of households who are 20 years old or over.

Table 1.—ANALYSIS OF SOCIAL SECURITY NUMBERS COLLECTED IN THE FIRST AND SECOND CONTENT PRETESTS, BY AGE

APPENDIX

Results of analysis	Percent—First and Second Content Pretests		Percent—First Content Pretest		Percent—Second Content Pretest	
	Total	Under 14 years old and over	Total	Under 14 years old and over	Total	Under 14 years old and over
A. Total persons in pre-test sample.....	(1)	(2)	(3)	(4)	(5)	(6)
B. Reported an SSN in pretest..	50.9	76.2	2.4	3.2	53.8	81.3
C. Correct SSN.....	44.5	66.6	2.1	3.2	46.9	70.8
D. Incorrect SSN.....	6.4	9.6	0.3	—	6.9	10.5
E. Reported not having an SSN	33.3	13.9	85.7	0.8	30.7	10.9
F. Possible SSN located.....	2.8	3.4	2.0	—	1.8	2.1
G. SSN not located.....	29.6	10.2	81.6	—	26.9	8.1
H. Insufficient information provided.....	0.9	0.3	2.1	0.8	2.0	0.8
I. Did not answer SSN item.....	15.9	9.9	11.9	96.0	15.5	7.8
J. Possible SSN located.....	3.2	4.8	0.2	—	2.7	4.1
K. SSN not located.....	6.6	5.0	11.7	1.2	4.3	3.7
L. Insufficient information provided.....	6.0	—	—	94.9	8.4	—
1/ Excluded from total since data for line L were not available in the Second Content Pretest.						

Table 2.—ANALYSIS OF SOCIAL SECURITY NUMBERS COLLECTED IN THE FIRST AND SECOND CONTENT PRETESTS, BY AGE
(Combined totals)

Age	Total persons in First and Second Content Pretests	Reported an SSN in pretest					Reported <u>not</u> having an SSN				Did not answer SSN item			
		Total	Correct SSN	Incorrect SSN			Total	Possible SSN located	SSN not located	Insufficient information provided	Total	Possible SSN located	SSN not located	Insufficient information provided
				Total	SSN located	SSN not located								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
A. Total.....	4,515	2,296	2,009	287	90	197	1,503	126	1,338	39	716	146	300	270
B. Under 14 years.....	1,270	31	27	4	2	2	1,088	25	1,036	27	151	3	148	—
C. 14 years and over....	2,962	2,257	1,974	283	88	195	413	101	302	10	292	143	149	—
D. 14 to 24 years.....	789	492	436	56	17	39	209	56	147	6	88	58	30	—
E. 25 to 44 years.....	921	774	679	95	30	65	56	12	41	3	91	44	47	—
F. 45 to 54 years.....	493	402	358	44	16	28	34	7	27	—	57	22	35	—
G. 55 to 61 years.....	300	239	210	29	11	18	39	11	28	—	22	7	15	—
H. 62 to 64 years.....	98	84	70	14	5	9	9	1	8	—	5	2	3	—
I. 65 to 71 years.....	192	144	124	20	4	16	36	5	31	—	12	4	8	—
J. 72 years and over..	169	122	97	25	5	20	30	9	20	1	17	6	11	—
K. Not reported.....	283	8	8	—	—	—	2	—	—	2	273	—	3	270
L. Total.....	100.0	50.9	44.5	6.4	2.0	4.4	33.3	2.8	29.6	0.9	15.9	3.2	6.6	6.0
M. Under 14 years.....	100.0	2.4	2.1	0.3	0.2	0.2	85.7	2.0	81.6	2.1	11.9	0.2	11.7	—
N. 14 years and over....	100.0	76.2	66.6	9.6	3.0	6.6	13.9	3.4	10.2	0.3	9.9	4.8	5.0	—
O. 14 to 24 years.....	100.0	62.4	55.3	7.1	2.2	4.9	26.5	7.1	18.6	0.8	11.2	7.4	3.8	—
P. 25 to 44 years.....	100.0	84.0	73.7	10.3	3.3	7.1	6.1	1.3	4.5	0.3	9.9	4.8	5.1	—
Q. 45 to 54 years.....	100.0	81.5	72.6	8.9	3.2	5.7	6.9	1.4	5.5	—	11.6	4.5	7.1	—
R. 55 to 61 years.....	100.0	79.7	70.0	9.7	3.7	6.0	13.0	3.7	9.3	—	7.3	2.3	5.0	—
S. 62 to 64 years.....	100.0	85.7	71.4	14.3	5.1	9.2	9.2	1.0	8.2	—	5.1	2.0	3.1	—
T. 65 to 71 years.....	100.0	75.0	64.6	10.4	2.1	8.3	18.7	2.6	16.1	—	6.2	2.1	4.2	—
U. 72 years and over..	100.0	72.2	57.4	14.8	3.0	11.8	17.8	5.3	11.8	0.6	10.1	3.6	6.5	—
V. Not reported.....	100.0	2.8	2.8	—	—	—	0.7	—	—	0.7	96.5	—	1.1	95.4

APPENDIX—Con.

Table 3.—ANALYSIS OF SOCIAL SECURITY NUMBERS COLLECTED IN THE FIRST AND SECOND CONTENT PRE-TESTS FOR PERSONS 14 YEARS OLD AND OVER, BY SEX

Results of analysis	Number		Percent	
	Males 14 years old and over Total	Females 14 years old and over Total	Males 14 years old and over Total	Females 14 years old and over Total
	(1)	(2)	(3)	(4)
A. Total persons in First and Second Content Pretest samples.....	1,386	1,576	100.0	100.0
B. Reported an SSN in pretest.....	1,150	1,107	83.0	70.2
C. Correct SSN.....	1,010	964	72.9	61.2
D. Incorrect SSN.....	140	143	10.1	9.1
E. Reported <u>not</u> having an SSN in pretest.....	96	317	6.9	20.1
F. Possible SSN located.....	29	72	2.1	4.6
G. SSN not located.....	64	238	4.6	15.1
H. Insufficient information.....	3	7	0.2	0.4
I. Did not answer SSN item.....	140	152	10.1	9.6
J. Possible SSN located.....	75	68	5.4	4.3
K. SSN not located.....	65	84	4.7	5.3
L. Insufficient information.....	—	—	—	—

Table 4.—ANALYSIS OF SOCIAL SECURITY NUMBERS COLLECTED IN THE SECOND CONTENT PRETEST FOR SPECIFIED PERSONS 14 YEARS OLD AND OVER

Results of analysis	Number		Percent	
	Heads of households Total	All other household members Total	Heads of households Total	All other household members Total
	(1)	(2)	(3)	(4)
A. Total persons in Second Content Pretest.....	723	932	100.0	100.0
B. Reported an SSN in pretest.....	620	574	85.8	61.6
C. Correct SSN.....	537	511	74.3	54.8
D. Incorrect SSN.....	83	63	11.5	6.8
E. Reported <u>not</u> having an SSN in pretest.....	36	235	5.0	25.2
F. Possible SSN located.....	11	64	1.5	6.9
G. SSN not located.....	25	171	3.5	18.3
H. Insufficient information.....	—	—	—	—
I. Did not answer SSN item.....	67	123	9.3	13.2
J. Possible SSN located.....	33	57	4.6	6.1
K. SSN not located.....	34	66	4.7	7.1
L. Insufficient information.....	—	—	—	—