Table of Contents

United States
Department of
Agriculture

Food and
Nutrition
Service

Food Stamp SSI/Elderly Cashout Demonstration Evaluation

Final Report

June 1982

Volume II: Methodological Appendices

Table of Contents

VOLUME II

METHODOLOGICAL APPENDICES

CONTENTS

Table of Contents

APPENDIX A:
WEIGHTING OF SURVEY DATA

The sample selection procedures (see description in Chapter VI) used for the eligibility/participation survey resulted in elderly households having different probabilities of selection into the sample, depending on a number of factors. This appendix outlines the algorithm used in computing weights to correct for this in the analysis tabulations.

The following factors affected selection probabilities:

Household size. The Master Beneficiary Record (MBR) sample frame included essentially all persons 65 years old and older in the study areas. As a result, households with more than one member had a greater probability of selection than did one-member households.

Sample frame. Households from the Supplemental Security Record (SSR) sample frame were oversampled in order to increase the efficiency with which the survey could be targeted on program-eligible persons.

Mail nonresponders. For the same efficiency reason, MBR sample members who responded to the mail survey were oversampled, as compared with nonresponders.

Phone/field. Households with locatable phone numbers were oversampled as compared with households without locatable phone numbers.

In order to account in the analysis tabulations for the unequal selection probabilities resulting from these factors, weighted tabulations were performed on the data, with greater weights being given to households with lower selection probabilities. In particular, each household was assigned a weight using the following algorithm:

$$W = WSIZE \times WFRAME \times WMAIL \times WPHONE,$$

where the factors on the right hand side of the equations are based on selection probabilities with regard to each of the sampling factors listed above.

The following sections describe how each of these factors was calculated.

WSIZE

In general, each household appeared only once on the SSR frame and therefore WSIZE = 1 for all SSR cases. For the MBR, each member of multiperson households was listed in the sample frame. Thus, if no attempt had been made to eliminate duplicate multiple members, the probability of selection for a two-person household would have been twice that for a one-person household, and similarly for larger sized households.

However, attempts were made to eliminate duplicates from the MBR frame, thus mitigating this effect somewhat. In particular, for each site, 6,000 persons were randomly chosen[1/] from the frame, and multiple household members were then

---

1/ Except for the South Carolina comparison site, where fewer than 6,000 sample points were available.

removed from this list of 6,000 names. Thus, for any site in which the entire frame consisted of more than 6,000 names, the probability of selection for a two-person household was still greater than that for a one-person household but was less than twice that of a one-person household. This was taken into account in computing WSIZE in the following way:

Because most elderly persons live in either one-person or two-person households, it was assumed as an approximation that all multi-person households contained two persons. Consider a two-person household, with members A and B. Without loss of generality, it can be assumed that A had the lower random number in the sample selection algorithm and was the member to be kept in the survey if it happened that both A and B were drawn into the 6,000 persons on the list from which duplicates had been eliminated. Then the probability of A being selected in the survey was

$$\frac{6,000}{P}$$

where P is the total number of sample frame persons for the site where A and B live. The probability of B being selected was the probability that B would be on the list of 6,000 names times the probability that A was not on this list. [If both were on the list, B was deleted.] This probability is given as:

$$\frac{6,000}{P}\left(1-\frac{6,000}{P}\right)$$

Thus, the combined probability that one member of the household, i.e., either A or B, was in the sample was:

$$\frac{6,000}{P}+\frac{6,000}{P}\left(1-\frac{6,000}{P}\right)$$

For a one-person household, the probability of selection was just:

$$\frac{6,000}{P}$$

Thus, the ratio of the probabilities was:

$$2-\frac{6,000}{P}$$

Therefore, WSIZE was set equal to 1 for a one-person household and set equal to:

$$\frac{1}{2-\frac{6,000}{P}}$$

for multi-person households.

The numbers of persons in the sample frame for each site were as follows:

	P	WSIZE
1. Monroe County, NY	75,717	0.5206
2. Albany County, NY	35,380	0.5463
3. Darlington and Dillon Counties, SC	8,600	0.7679
4. Lee and Marlboro Counties, SC	4,833	1.0000
5. Multnomah County, OR	86,296	0.5180
6. Lane County, OR	25,013	0.5681

WFRAME

Across all sites, the total size of the SSR sample frame was 9,850 persons with positive SSI payments. The total number of SSR cases released into the phone/field survey work was 5,608. Thus, the probability of selection into the SSR sample was .58. The total size of the MBR frame was 235,839 persons, and the total number of persons drawn into the lists was 33,180. Therefore, the probability of selection for MBR sample members can be approximated as .14. Thus, the probability of selection for SSR cases was approximately 4.1 times that of MBR cases. To correct for this, WFRAME was set equal to 1 for SSR cases and equal to 4.1 for MBR cases.

WMAIL

There was no mail prescreening for the SSR sample frame, and therefore WMAIL was set at 1 for SSR sample frame members. For MBR sample members, all cases that returned the mail survey were tracked into the phone/field survey, if their mail survey responses indicated they were eligible. Only 6,192 mail survey nonresponders, out of a total of 12,740, were tracked into the phone/field survey. Thus, compared with responders, nonresponders had a selection probability only 6,192/12,740 as large, and WMAIL was set at 12,740/6,192 or 2.1 for mail nonresponders.

WPHONE

In the initial sample frames, approximately 70 percent of households had locatable phone numbers.[1/] Within the samples actually released into the phone and field surveys, approximately 84 percent had phone numbers. Therefore, the probability of selection for households with locatable phone numbers was:

[1/] The MBR sample frame consisted of 235,839 persons. Assuming as an approximation an average of 1.46 persons per household (based on preliminary tabulations of the survey data), there were an estimated 161,534 households. The response rate to the mail survey was .48 and the rate at which responders had locatable phone numbers was approximately .75. Thus 161,534 times .48 times .75, or 58,152 households were potential mail responders and had phone numbers. Similarly, the rate of nonresponse was .52 and the rate of locatable phone numbers for this nonresponder group was .82, so that the total number of households from this group with locatable phone numbers was 52,079. Overall, therefore, the MBR universe included approximately 110,230 households with locatable phones. Among the 12,194-household SSR universe, approximately .88 had locatable phone numbers. Thus, there were approximately 10,730 SSR households with phone numbers. Overall, across both sample frames, therefore, approximately 120,960 of 173,728 households had phone numbers, or about 70 percent.

$$\frac{.34}{.70}\left[\frac{S}{U}\right] = 1.2 \left[\frac{S}{U}\right]$$

where U is the total universe of households and S is the total sample size for the phone/field survey. Similarly, the probability of selection for households without locatable phone numbers was:

$$\frac{.16}{.30}\left[\frac{S}{U}\right] = .53 \left[\frac{S}{U}\right]$$

Therefore, the probability of being selected was 2.3 times as great for households with locatable phone numbers than for those without. To correct for this, WPHONE was set at 1 for households in the phone sample and at 2.3 for households in the field sample.

Aggregation Across Sites

Parts of the analysis involve tabulations aggregated across sites. No differential weights by site were used for this work because the number of sites at which it was feasible to conduct the survey was too small to permit statistically rigorous generalization to a national universe. In selecting the sites, an effort was made to choose those representative of the country at large, and it is reasonable to hope that the data obtained reflect national conditions. From a rigorous statistical point of view, however, reliable national generalizations cannot be made. Therefore, there was no basis for developing weights to produce such generalizations.

Table of Contents

APPENDIX B: SURVEY NONRESPONSE ANALYSIS

As discussed in more detail in Volume III of the report, which describes the data collection for the project, the survey work met with considerable nonresponse. The estimated response rate in the combined phone/field interviewing for the eligibility/participation survey was approximately 65 percent. Table B.1, which is discussed more fully in Volume III, summarizes reasons for nonresponse.

Given the level of nonresponse, it is of interest to examine evidence concerning whether the respondents to the survey were similar to the sample members who did not complete the interview. The sample frames from which the samples were drawn include data that can be used for this purpose. Both the Master Beneficiary Record (MBR) sample frame and the Supplemental Security Record (SSR) frame include the dates of birth of the sample members. In addition, the MBR frame includes monthly Social Security payment data, while the SSR frame has comparable information regarding SSI payments.

Table B.2 presents these data for respondents and nonrespondents to the eligibility/participation survey. MBR sample respondents were, on average, about half a year younger than nonrespondents and their Social Security payments were $19 lower. Because of the very large sample sizes available for this work, these differences are statistically significant.

SSR sample respondents were a year younger and their SSI payments were $2 less. The difference in age is statistically significant but the differences in SSI receipts is not.

These results show that there do appear to be some systematic differences between respondents and nonrespondents. The implications these differences have for the analysis are discussed in Chapter VI of the report.

Table of Contents

TABLE B.1

INTERVIEW STATUS BY INTERVIEW METHOD

	Phone	Field	Total
b. Household Found to Include Members Under 65	2578	552	3130
c. Institutionalized	328	232	560
d. Moved Out of Area	34	45	79
e. Deceased	256	104	360
f. Not Located	1197	258	1455
g. Refused	3277	349	3626
h. Non-English Speaking	193	21	214
i. Physically Impaired	359	41	400
j. Unable to Contact	299	79	378
TOTAL SAMPLE	12,703	2,409	15,112
Eligible for Interviewing[a/]	11,218	2,002	13,218
RESPONSE RATE [b/]	63.2	75.5	65.1

[a/] Calculated by deducting sample members who were deceased, not located, or moved out of the area from the total sample.

[b/] Response rate is defined as the percentage of the sample members eligible for interviewing for whom cashout demonstration eligibility was determined (a, b, and c above).

Table of Contents

TABLE B.2

COMPARISON OF RESPONDENTS AND NONRESPONDENTS IN ELIGIBILITY/PARTICIPATION SURVEY

	Respondents [a]	Nonrespondents	Difference
MBR Sample			
Age (years)	75.12 [b]	75.66	-.54
	(.09)	(.16)	(.18)
Monthly Social Security Payment	$325	$344	-$19
	(2)	(2)	(3)
SSR Sample			
Age (years)	76.00	77.00	-1.00
	(.13)	(.26)	(.32)
Monthly SSI payment	$111	$113	-$2
	(2)	(3)	(4)

[a] Respondents and nonrespondents are defined in footnote to Table B.1.

[b] Standard errors of estimates appear in parentheses under table entries.

APPENDIX C: SAMPLING ERROR IN TABULAR ESTIMATES

The sample stratification described in Appendix A increases the sampling errors of estimates based on survey data tabulations beyond what they would be if a simple random sample of the same size had been used. Overall sampling errors for proportions estimated in Chapter VII can be estimated as

$$\text{Standard Deviation} = \sqrt{(\text{d.e.}) \times \frac{(p)(1-p)}{n}}$$

where d.e. is the design effect resulting from the stratification, and the remainder of the equation is based on the standard estimator for the variance of the estimated mean of a binomial distribution.

Tables C.1 and C.2 present approximate standard errors for various proportion estimates and sample sizes, based on the above equation. Design effects have been estimated using the following equation:

$$\text{d.e.} = \frac{(\Sigma W^2)(\text{sample size})}{(\Sigma W)^2}$$

where the W's are the weights described in Appendix A. [The formula is derived from Cochran (1977), p. 92, taking into account that the weights in the current survey have not been normalized to add to the sample size.]

The estimated design effects are 2.77 for the participant sample and 1.85 for the nonparticipant sample.

Table of Contents

TABLE C.1

APPROXIMATE STANDARD ERRORS FOR PROPORTION ESTIMATES BASED ON TABULATIONS OF PARTICIPANT SURVEY DATA

Proportion Estimate	Sample Size		
	50	200	800
.1	.07	.04	.02
.3	.11	.05	.03
.5	.12	.06	.03
.7	.11	.05	.03
.9	.07	.04	.02

Table of Contents

TABLE C.2

APPROXIMATE STANDARD ERRORS FOR
PROPORTION ESTIMATES BASED ON TABULATIONS
OF NONPARTICIPANT SURVEY DATA

Proportion Estimate	Sample Size		
	50	200	800
.1	.06	.03	.01
.3	.09	.04	.02
.5	.10	.05	.02
.7	.09	.04	.02
.9	.06	.03	.01

Table of Contents

WRONG FONT SIZE

APPENDIX D:
SAMPLES FOR ANALYSIS OF RANDOM CASE RECORDS

This appendix describes the samples used for the descriptive analysis of case records data presented in Chapter IV. For the three sites at which survey operations were conducted, New York, Oregon, and South Carolina, machine-readable case records data were available on all program participants. Therefore, the samples used in the descriptive case records analysis for these sites consisted of all participant households. The sample used for the Virginia site, which had the least number of participants, also consisted of all the participant households. For the remaining four sites, random samples of approximately 500 to 600 households were drawn, and key case records variables were manually coded from case records data supplied by the sites. Table D.1 shows the sample sizes for each site. All of the samples were drawn during the second half of the planned one-year demonstration evaluation period.

Because the tabulations for the three survey sites and for Virginia are based on all cases, these data involve no sampling error. For the other four sites, the estimates presented in the text of Chapter IV are subject to some degree of sampling error. The approximate sizes of such errors are given in Table D.2, which shows the width of 95 percent confidence intervals associated with percentage estimates based on a sample of 500 cases. It should be noted that this table provides an upper bound of sampling error, particularly for the smaller sites, because for simplicity it ignores reductions in variance estimates due to finite sample size corrections.

In performing the data tabulations, cases with missing data were omitted. For most data items, cases with missing data accounted for fewer than 10 percent of all cases. The only significant exception, as indicated in Table IV.2 in the text of the report, is that certain data items—most frequently gross income—were entirely unavailable in certain sites.

All tabulations were weighted, with each observation having a weight equal to the inverse of its probability of selection. The weights were based on the numbers of participating households shown in Table III.2 of Volume I of the report and the sample sizes in Table D.1.

Table of Contents

TABLE D.1

CASE RECORDS SAMPLES

Site	# Cases in Sample	Date of Sample
Utah	580	01/06/81
South Carolina	3,659	02/22/81
Oregon	5,826	04/27/81
Hennepin County, MN	567	03/27/81
Monroe County, NY	4,128	03/04/81
Vermont	549	03/23/81
Cuyahoga County, OH	500	03/13/81
Virginia	477	01/31/81

Table of Contents

TABLE D.2

SAMPLING ERROR IN ESTIMATING PERCENTAGES
USING 500 OBSERVATIONS

True Percentage	Width of 95 Percent Confidence Interval
10	± .026
30	± .040
50	± .044
70	± .040
90	± .026

Table of Contents

APPENDIX E:
SURVEY DATA BY SITE

This appendix presents survey data tabulated by site. All tabulations are weighted as described in Appendix A. Numbering of tables corresponds to table numbers in the text of the report. For instance, Table E.VII.1 presents site-by-site data for the variables included in Table VII.1 of the report.

Table of Contents

TABLE E.VI.1

CHARACTERISTICS OF SAMPLE

	Participants						NonParticipants					
	NY Dem	NY Comp	SC Dem	SC Comp	OR Dem	OR Comp	NY Dem	NY Comp	SC Dem	SC Comp	OR Dem	OR Comp
Household Size												
1	96	91	69	72	93	95	86	88	71	71	89	84
≥ 2	4	9	31	28	7	5	14	12	29	29	11	16
Sex of Head												
Male	30	26	30	40	26	17	23	34	44	38	29	28
Female	70	74	70	60	74	83	77	66	56	62	71	72
Age of Head												
65 - 69	34	29	31	29	34	31	20	19	28	33	19	18
70 - 74	29	38	34	31	30	24	23	39	31	31	30	31
75 - 79	20	16	21	20	16	21	25	16	31	14	28	25
≥ 80	17	17	14	19	20	24	30	25	10	22	23	26
Race of Head												
Black	22	23	57	65	10	0	7	10	45	49	9	0
White	78	76	43	35	89	99	93	89	55	51	90	100
Other	0	1	0	0	1	1	0	1	0	0	1	0
Education of Head												
0 - 8 years	72	68	81	90	46	39	56	54	69	79	48	50
9 - 11 years	18	21	15	6	23	27	18	25	18	9	13	20
≥ 12 years	10	11	4	4	31	34	26	21	13	12	39	30
Monthly Income												
$0 - 100	0	0	3	1	0	1	0	0	0	3	0	0
101 - 200	1	1	4	9	6	2	0	2	15	15	3	1
201 - 300	8	15	44	47	54	53	7	9	31	17	33	27
301 - 400	84	70	32	16	27	29	49	54	26	29	41	46
401 - 500	4	5	12	19	12	13	35	12	12	26	11	10
501 - 600	1	8	2	6	1	2	9	16	10	7	7	9
601 - 700	2	0	2	2	0	0	0	1	5	3	5	6
701 - 800	0	0	1	0	0	0	0	3	1	0	0	1
> 800	0	0	0	0	0	0	0	3	0	0	0	0
Sources of Income												
Social Security	85	90	90	90	92	92	97	96	98	97	93	98
SSI	64	71	49	61	53	52	17	26	12	14	18	14
Earnings	0	1	2	3	2	1	0	4	11	11	2	2
Pensions	15	6	9	5	10	11	13	10	14	23	21	26
Other	7	7	5	3	21	13	13	18	11	12	5	15
Sample Size[a]	194	181	326	328	234	232	72	143	194	145	103	141

[a] Individual item tabulations may be based on smaller sample sizes because cases with missing data were excluded.

NB: Numbers are percentages (except sample sizes).

Table of Contents

TABLE E.VI.4

RESPONSES TO QUESTIONS RELATING TO STIGMA

	Percentage											
	Participants						NonParticipants					
	NY Dem	NY Comp	SC Dem	SC Comp	OR Dem	OR Comp	NY Dem	NY Comp	SC Dem	SC Comp	OR Dem	OR Comp
1. "Bothered" by receiving food stamps												
Yes	24	16	17	19	25	30	38	32	15	23	48	46
No	76	84	83	81	75	70	62	68	85	77	52	54
2. Degree of embarrassment at telling friends they receive food stamps												
"very embarrassed"	9	6	1	3	12	13	19	22	3	15	20	25
"somewhat embarrassed"	19	12	4	18	14	16	19	21	7	17	23	24
"not embarrassed at all"	72	82	95	79	74	71	62	57	90	68	57	51
3. Perceive people in community as having less respect for food stamp recipients												
Yes	17	7	15	24	22	17	43	17	18	31	16	15
No	59	61	65	68	54	63	41	48	60	48	53	54
Don't know	24	32	20	8	24	20	16	35	22	21	31	31

Table of Contents

TABLE E.VI.5

RESPONSES TO QUESTIONS RELATING TO
FOOD STAMP OFFICE ACCESS

	Percentage											
	Participants						NonParticipants					
	NY Dem	NY Comp	SC Dem	SC Comp	OR Dem	OR Comp	NY Dem	NY Comp	SC Dem	SC Comp	OR Dem	OR Comp
Perceive getting to program office as a problem												
"big problem"	33	25	26	38	25	19	52	29	20	34	29	32
"little problem"	14	23	31	36	21	25	8	27	26	31	24	28
"no problem"	53	52	43	26	54	56	40	44	54	35	47	40
Distance to FS office												
< 1 mi	13	46	18	4	18	20	11	20	11	2	10	8
1-2 mi	28	32	37	23	42	28	17	40	40	33	45	29
2-4 mi	14	8	11	19	20	16	7	16	10	10	13	13
4-9 mi	16	6	22	12	16	19	48	12	16	19	21	9
> 9 mi	29	8	12	41	4	17	16	12	23	36	11	40
Own car	12	7	29	24	12	35	24	14	46	41	26	43
Own or have access to car	62	51	75	71	59	76	70	65	89	91	73	82

Table of Contents

TABLE E.VI.6

MONTHLY FOOD STAMP ENTITLEMENTS

	Percentage											
	Participants						NonParticipants					
	NY Dem	NY Comp	SC Dem	SC Comp	OR Dem	OR Comp	NY Dem	NY Comp	SC Dem	SC Comp	OR Dem	OR Comp
$10 - 15	63	36	14	19	30	21	55	49	29	36	40	47
16 - 30	9	12	22	34	26	23	11	18	21	14	19	16
31 - 45	7	22	24	18	17	29	14	13	11	17	17	16
46 - 60	8	13	17	13	13	10	2	7	12	7	5	4
61 - 75	12	17	19	13	14	13	16	12	24	28	19	16
76 - 90	0	0	2	1	0	2	0	1	3	0	0	1
More than 90	1	0	2	2	0	2	0	0	0	0	0	0

Table of Contents

TABLE E.VIII.1

TABULATION OF AWARENESS OF AND ATTITUDES TOWARDS CASHOUT

	Percentage											
	Participants						NonParticipants					
	NY Dem	NY Comp	SC Dem	SC Comp	OR Dem	OR Comp	NY Dem	NY Comp	SC Dem	SC Comp	OR Dem	OR Comp
1. Had heard of cashout program	NA	NA	NA	NA	NA	NA	48	8	61	16	39	18
2. Attitudes toward cashout												
Prefer checks	77	29	74	36	80	29	49	48	53	50	36	45
Prefer coupons	5	26	6	20	9	34	17	22	11	9	19	10
No Opinion	18	45	20	44	11	37	34	30	36	41	45	45

Table of Contents

TABLE E.VIII.2

REASONS FOR PREFERRING CHECKS

	Percentage											
	Participants						NonParticipants					
	NY Dem	NY Comp	SC Dem	SC Comp	OR Dem	OR Comp	NY Dem	NY Comp	SC Dem	SC Comp	OR Dem	OR Comp
Checks more convenient or easier to use	87	69	85	76	70	57	60	73	71	58	69	56
Checks can be used for anything	30	20	27	36	20	35	30	10	39	40	10	31
Stamps inconvenient	7	5	9	8	7	22	0	4	1	2	1	0
With checks people don't know you get food stamp benefits or with checks you feel more dignified, not embarrassed	22	25	5	40	21	27	20	27	14	40	38	45

Table of Contents

TABLE E.VIII.3

REASONS FOR PREFERRING STAMPS

	Percentage											
	Participants						NonParticipants					
	NY Dem	NY Comp	SC Dem	SC Comp	OR Dem	OR Comp	NY Dem	NY Comp	SC Dem	SC Comp	OR Dem	OR Comp
Stamps more convenient	51	40	33	34	21	33	14	46	17	42	9	3
Checks difficult to cash	0	3	0	3	0	0	11	2	5	5	0	0
Stamps ensure food stamp benefits are spent for food	20	49	37	41	45	60	86	67	42	72	78	92
Other	51	11	41	25	39	9	0	1	18	15	14	2

Table of Contents

TABLE E.VI.2

PAST PROGRAM EXPERIENCE OF ELIGIBLE NONPARTICIPANTS

	Percentage					
	NY Dem	NY Comp	SC Dem	SC Comp	OR Dem	OR Comp
1. Percentage who tried to determine eligibility for food stamps	37	40	57	46	35	45
2. Percentage who applied for food stamps	21	36	43	43	29	28
3. Percentage who received food stamps	21	23	26	26	19	19
4. Disposition of application for those who applied but never received food stamps						
Application denied	100	81	68	91	90	82
Changed mind; chose to do without	0	2	11	0	8	8
Other	0	17	23	9	2	10
5. Reason given for termination of food stamp benefits by those who at one time received them						
- Family began earning too much money	37	28	33	35	20	33
- Recertification took too long	0	20	8	1	3	17
- Inconvenient	0	3	3	10	42	10
- Tranportation problem	10	24	14	12	10	9
- Food stamps cost too much	2	10	5	24	5	13
- Other	51	15	37	18	20	18
6. Percentage who believe themselves eligible for food stamps						
Believe eligible	25	45	42	21	31	23
Believe ineligible	42	31	34	24	42	38
Don't know	33	24	24	55	27	39

Table of Contents

TABLE E.VI.3

STATED REASONS FOR NONPARTICIPATION

	Nonparticipants Who Never Applied (Percentage)					
	NY Dem	NY Comp	SC Dem	SC Comp	OR Dem	OR Comp
Believe ineligible	16	34	40	20	15	17
Don't need the benefits	48	46	26	30	44	42
The benefits don't seem worth the trouble	22	13	17	34	13	19
Would be embarrassed if other people knew	7	0	1	9	1	7
Don't know how to apply	5	0	7	0	0	2
Couldn't get to the office	4	2	1	8	0	4
Too proud to apply	8	11	3	10	20	14
Stamps cost too much	0	0	3	2	1	1
Never thought about it	7	10	12	22	13	5

Table of Contents

TABLE E.VI.7

INTERVIEW RESPONSES RELATED TO PERCEPTION OF FOOD STAMP PROGRAM

	Percentage					
	NY Dem	NY Comp	SC Dem	SC Comp	OR Dem	OR Comp
1. Perception of experiences at food stamp office by nonparticipants who had applied						
a. How treated						
"treatment was fine"	94	91	79	87	96	71
"people were rude"	6	3	5	17	4	17
b. Helpfulness of program staff						
"people were helpful"	58	83	25	47	79	75
"people were not helpful"	42	17	75	53	21	25
2. Perception by participants of "What kind of job Food Stamp Program is doing to take care of their food needs"						
Good	59	69	24	39	39	42
Fair	41	21	34	37	27	37
Poor	0	10	41	23	23	20

Table of Contents

TABLE E.VII.1

PERCEIVED EFFECTS OF FOOD STAMP BENEFITS
ON FOOD BUYING

	Percentage					
	NY Dem	NY Comp	SC Dem	SC Comp	OR Dem	OR Comp
1. Effect on Amount of Food						
More	58	56	57	63	59	62
Less	1	0	4	3	2	1
Same	41	44	35	33	37	37
2. Effect on Quality of Food						
Better	30	35	32	31	37	37
Lower	4	1	1	1	2	2
Same	66	64	64	66	61	58
3. Percentage Reporting an Increase in Either Quantity or Quality	48	58	56	61	59	67
4. Percentage Reporting a Decrease in Either Quantity or Quality	3	0	4	4	3	2

Table of Contents

TABLE E.VII.2

PERCEIVED EFFECTS OF SWITCHING
FROM FOOD STAMPS TO CHECKS

	Percentage		
	NY Dem	SC Dem	OR Dem
1. Effect on Amount of Food			
More	6	8	5
Less	9	12	16
Same	84	78	74
2. Effect on Quality of Food			
Better	6	3	5
Lower	3	6	7
Same	91	86	81
3. Percentage Reporting an Increase in Either Quantity or Quality	7	8	7
4. Percentage Reporting a Decrease in Either Quantity or Quality	8	12	17

Table of Contents

TABLE E.VIII.4

CLIENT EXPERIENCE WITH CHECKS

	Percentage		
	NY Dem	SC Dem	OR Dem
Percentage reporting checks arriving late	8	60	24
Percentage reporting checks stolen	5	1	4
Percentage reporting check cashing fee	3	2	1
Median check cashing fee among those reporting fee	$.50	$.50	$.50

Table of Contents

APPENDIX G: DATA SET USED TO ANALYZE CHANGES IN REPORTED MONTHLY PARTICIPATION

This appendix describes the data used in the analysis of changes in monthly participation reported in Chapter V. Table G.1 describes all the data available when the analysis was conducted. Table G.2 describes the reports used in the analysis, and Table G.3 lists the data used. Tables G.4 through G.7 display estimated changes in participation.

In general, the analysis was conducted using the first twelve months of data available for each site. However, several exceptions should be noted. For several sites, fewer than twelve months of data were available, and for those sites, the longest available data set was used. For the Albany, New York site, only five months of monthly report data were available. However, a computer listing of participants as of September 1981 had been obtained as part of the survey work, and for that site, counts of that listing were used as the end-of-period data for the analysis.

The comparison and supplemental sites in South Carolina and Wyoming indicated very large changes in participation in all groups during the first few months of the demonstration, and the comparison site in Minnesota did so in the non-SSI aged category. None of these changes was plausible in terms of known events, and in several instances the data were inconsistent with more detailed case records data supplied by the sites. Also, the Minnesota reports were inconsistent with reports of new participation, which appeared to be reasonable in that site. Therefore, in the South Carolina comparison and supplemental sites and in Wyoming, the first three months of data were not used in the analysis. In the Minnesota comparison site, the total participation in the non-SSI aged category was adjusted in the following way: data supplied by the site showed that over the period in question, there were 160 new households in the non-SSI aged category. The reported number of households in this category at the end of the period was 332, and this was assumed to be accurate. Also, it was assumed on the basis of data from the non-SSI aged category for the site that approximately 41 percent of the caseload left the program during the period. The net change in the non-SSI aged category was then estimated as [160 - .41 x 332] = 24. This change was used to estimate the non-SSI aged caseload for the beginning of the period.

One county in South Carolina apparently reversed the non-SSI aged and SSI aged columns on the forms it submitted. The reversal was confirmed by examining case records data and then corrected by changing the forms before analyzing them.

Table of Contents

TABLE G.1

AVAILABLE REPORTS OF TOTAL PARTICIPATION IN EACH SITE

Sites	Largest City	First Month	Last Month	# Months Reported	Missing
Vermont (D)[a/]		07/80	11/81	15	10/81
Clinton Co., NY (C)	Plattsburg	07/80	04/81	9	
Essex Co., NY (C)	Saranac Lake	08/80	12/80	4	09/80
Hennepin Co., MN (D)	Minneapolis	05/80	10/81	18	
St. Louis Co., MN (D)	Duluth	05/80	07/81	15	
Marion Co., IN (S)	Indianapolis	05/80	03/81	11	
Arlington, VA (D)		09/80	08/81	12	
Alexandria, VA (C)		09/80	08/81	12	
Two Regions of OR (D)	Portland	08/80	09/81	14	
Lane Co., OR (C)	Eugene	08/80	09/81	14	
Balance of State, OR (S)		08/80	09/81	14	
Monroe Co., NY (D)	Rochester	06/80	09/81		12/80-05/81
Albany Co., NY (C)	Albany	06/80	12/80	5	07/80-08/80
Erie Co., NY (C)	Buffalo	07/80	02/81	8	
Four Counties of SC (D)	Florence	04/80	06/81	14	03/81
Three Counties of SC (C)	Orangeburg	04/80	03/81	10	Marlboro-04/80 Lee 3/81 Orangeburg-11/80,2/81
Lancaster Co., SC (S)	Lancaster	04/80	03/81	12	
Cuyahoga Co., OH (D)	Cleveland	05/80	09/81	17	
Franklin Co., OH (C)	Columbus	05/80	04/81	12	
Hamilton Co., OH (S)	Cincinnati	05/80	04/81	12	
Utah (D)		04/80	10/81	14	03,05,06,08,09/80
Wyoming (C)		04/80	02/81	11	
Tulsa Co., OK (S)	Tulsa	04/80	03/81	12	

[a/] D=Demonstration; C=Comparison; S=Supplemental.

Table of Contents

TABLE G.2

REPORTS USED IN THE ANALYSIS OF TOTAL PARTICIPATION

Sites	Largest City	First Month	Last Month	Length of Period (Mos)
Vermont (D)[a/]		07/80	07/81	12
Clinton Co., NY (C)	Plattsburg	07/80	04/81	9
Essex Co., NY (C)	Saranac Lake	08/80	12/80	4
Hennepin Co., MN (D)	Minneapolis	05/80	05/81	12
St. Louis Co., MN (C)	Duluth	05/80	05/81	12
Marion Co., IN (S)	Indianapolis	05/80	03/81	10
Arlington, VA (D)		09/80	09/81	12
Alexandria, VA (C)		09/80	08/81	11
Two Regions of OR (D)	Portland	08/80	08/81	12
Lane Co., OR (C)	Eugene	08/80	08/81	12
Balance of State, OR (S)		08/80	08/81	12
Monroe Co., NY (D)	Rochester	06/80	06/81	12
Albany Co., NY (C)	Albany	06/80	09/81	15
Erie Co., NY (C)	Buffalo	07/80	02/81	7
Four Counties of SC (D)	Florence	04/80	04/81	12
Three Counties of SC (C)	Orangeburg	07/80	03/81	8
Lancaster Co., SC (S)	Lancaster	04/80	03/81	8
Cuyahoga Co., OH (D)	Cleveland	06/80	05/81	12
Franklin Co., OH (C)	Columbus	05/80	04/81	11
Hamilton Co., OH (S)	Cincinnati	05/80	04/81	11
Utah (D)		04/80	04/81	12
Wyoming (C)		07/80	02/81	7
Tulsa Co., OK (S)	Tulsa	04/80	03/81	11

NOTE: See text of appendix for criteria used in selecting analysis periods.

[a/] D=Demonstration; C=Comparison; S=Supplemental.

Table of Contents

TABLE G.3

DATA USED IN ANALYSIS OF TOTAL PARTICIPATION

Sites	Beginning of Period				End of Period			
	Aged	SSIA[a]	SSIBD[a]	Total	Aged	SSIA[a]	SSIBD[a]	Total
Vermont (D)[b]	1,264	1,576	1,060	3,900	1,528	1,763	1,155	4,446
Clinton and Essex Counties, NY (C)	171	553	250	974	220	538	374	1,132
Hennepin Co., MN (D)	1,058	945	1,004	3,007	1,146	1,112	1,263	3,521
St. Louis Co., MN (C)	308	398	222	928	332	382	213	927
Marion Co., IN (S)	908	920	453	2,281	856	901	412	2,169
Arlington, VA (D)	170	161	121	452	188	193	131	512
Alexandria, VA (C)	195	184	193	572	175	182	195	552
Two Regions of OR (D)	1,933	1,729	1,928	5,590	1,989	1,688	2,043	5,718
Lane Co., OR (C)	820	886	730	2,436	864	966	876	2,700
Balance of State, OR (S)	1,877	1,501	1,318	4,696	1,901	1,628	1,485	5,010
Monroe Co., NY (D)	541	1,451	1,803	3,795	634	1,493	1,869	3,996
Albany County and Erie Counties, NY (C)	2,405	3,826	4,957	11,188	2,624	4,218	5,876	12,518
Four Counties of SC (D)	617	1,788	917	3,322	752	1,941	1,040	3,733
Three Counties of SC (C)	531	1,252	542	2,325	549	1,306	591	2,446
Lancaster Co., SC (S)	176	288	68	532	185	291	80	556
Cuyahoga Co., OH (D)	3,147	3,774	4,598	11,519	3,308	3,607	5,153	12,068
Franklin Co., OH (C)	1,084	1,377	2,431	4,892	967	1,417	2,479	4,863
Hamilton Co., OH (S)	893	1,939	2,069	4,901	709	1,581	1,961	4,251
Utah (D)	996	1,215	1,076	3,287	1,161	1,169	1,186	3,516
Wyoming (C)	348	200	204	752	371	255	245	871
Tulsa Co., OK (S)	1,017	1,441	629	3,087	964	1,451	838	3,053
Total of the Eight Demonstration Sites	9,726	12,839	12,507	34,872	10,706	12,964	13,840	37,510
Total of the Other Sites	10,733	14,765	14,068	39,564	10,717	15,116	15,225	41,048

[a] SSIA = SSI Aged; SSIBD = SSI Blind and Disabled.

[b] D=Demonstration; C=Comperison; S=Supplemental.

Table of Contents

TABLE G.4

CHANGES IN TOTAL CASELOADS: ALL CATEGORIES

Site (State)	(1) Demonstration Beginning	(1) Demonstration End	(1) Demonstration % Change	(2) Comparison Beginning	(2) Comparison End	(2) Comparison % Change	(3) Comparison & Supplemental Beginning	(3) Comparison & Supplemental End	(3) Comparison & Supplemental % Change	Differences (1)-(2)	Differences (1)-(3)
Vermont	3900	4446	14.0%	974	1132	16.2%	974	1132	6.2%	-2.2%	-2.2%
Minnesota	3007	3521	17.1	928	927	-0.1	3208	3096	-3.5	17.2	20.6
Virginia	452	512	13.3	572	552	-3.5	572	552	-3.5	16.8	16.8
Oregon	5590	5718	2.3	2436	2700	10.8	7132	7710	8.1	-8.5	-5.8
New York	3795	3996	5.3	11188	12518	11.9	11188	12518	11.9	-6.6	-6.6
South Carolina	3322	3733	12.4	2325	2446	5.2	2857	3002	5.1	7.2	7.3
Oregon	11519	12068	4.8	4892	4986	-0.6	9793	9114	-6.9	5.4	11.7
Utah	3287	3516	7.0	752	871	15.8	3839	3924	2.2	-8.9	4.8
Unweighted Averages			9.5			7.0			3.7	2.5	5.8
t-values										(0.7)	(1.6)

TABLE 6.5

CHANGES IN TOTAL CASELOADS: NON-SSI AGED

Site (State)	(1) Demonstration Beginning	(1) Demonstration End	(1) Demonstration % Change	(2) Comparison Beginning	(2) Comparison End	(2) Comparison % Change	(3) Comparison & Supplemental Beginning	(3) Comparison & Supplemental End	(3) Comparison & Supplemental % Change	Differences (1)-(2)	Differences (1)-(3)
Vermont	1264	1528	20.9%	171	220	28.7%	171	220	28.7%	-7.8%	-7.8%
Minnesota	1058	1146	8.3	308	332	7.8	1216	1188	-2.3	0.5	10.6
Virginia	170	188	10.6	195	175	-10.3	195	175	-10.3	20.8	20.8
Oregon	1933	1989	2.8	820	864	5.4	2687	2765	2.5	-2.5	0.4
New York	541	634	17.2	2405	2624	9.1	2405	2624	9.1	8.1	8.1
South Carolina	617	752	21.9	531	549	3.4	707	734	3.8	18.5	18.1
Oregon	3147	3308	5.1	1084	967	-10.8	1977	1676	-15.2	15.9	20.3
Utah	996	1161	16.6	348	371	6.6	1365	1335	-2.2	10.0	18.8
Unweighted Averages			12.9			5.0			1.8	7.9	11.2
t-values										(2.2)	(3.0)

TABLE 6.6

CHANGES IN TOTAL CASELOADS: SSI AIED

Site (State)	(1) Demonstration Beginning	End	% Change	(2) Comparison Beginning	End	% Change	(3) Comparison & Supplemental Beginning	End	% Change	Differences (1)-(2)	(1)-(3)
Vermont	1576	1763	11.9%	553	538	-2.7%	553	538	-2.7%	14.6%	14.6%
Minnesota	945	1112	17.7	398	382	-4.0	1318	1283	-2.7	21.7	20.3
Virginia	161	193	19.9	184	182	-1.1	184	182	-1.1	21.0	21.0
Oregon	1729	1686	-2.5	886	966	9.0	2387	2594	8.7	-11.5	-11.2
New York	1451	1493	2.9	3826	4218	10.2	3826	4218	10.2	-7.4	-7.4
South Carolina	1788	1941	8.6	1252	1306	4.3	1540	1597	3.7	4.2	4.9
Oregon	3774	3607	-4.4	1377	1417	2.9	3316	2998	-9.6	-7.3	5.2
Utah	1215	1169	-3.8	200	255	27.5	1641	1706	4.0	-31.3	-7.7
Unweighted Averages			6.3			5.8			1.3	0.5	5.0
t-values										(0.1)	(1.1)

Table of Contents

can't find identifier

TABLE G.7

CHANGES IN TOTAL CASELOADS: SSI BLIND AND DISABLED

oops!

Site (State)	(1) Demonstration Beginning	(1) Demonstration End	(1) Demonstration % Change	(2) Comparison Beginning	(2) Comparison End	(2) Comparison % Change	(3) Comparison & Supplemental Beginning	(3) Comparison & Supplemental End	(3) Comparison & Supplemental % Change	Differences (1)-(2)	Differences (1)-(3)
Vermont	1060	1155	9.0%	250	374	49.6%	250	374	49.6%	-40.6%	-40.6%
Minnesota	1004	1263	25.8	222	213	-4.1	675	625	-7.4	29.9	33.2
Virginia	121	131	8.3	193	195	1.0	193	195	1.0	7.2	7.2
Oregon	1928	2043	6.0	730	876	20.0	2048	2361	15.3	-14.0	-9.3
New York	1803	1869	3.7	4957	5676	14.5	4957	5676	1.45	-10.8	-10.8
South Carolina	917	1040	13.4	542	591	9.0	610	671	10.0	4.4	3.4
Oregon	4598	5153	12.1	2431	2479	2.0	4500	4440	-1.3	10.1	13.4
Utah	1076	1186	10.2	204	245	20.1	833	883	6.0	-9.9	4.2
Unweighted Averages			11.1			14.0			11.0	-3.0	0.1
t-values										(-0.4)	(0.1)

Table of Contents

Row 5 is the sum of Rows 3 and 4.

Rows 6 and 7 are computed by dividing Rows 3 and 4 by Rows 1 and 2, respectively.

Row 8 is the sum of Rows 6 and 7.

Row 9 is Row 5 divided by Row 8. Standard errors have been computed in the following way:

The entries in Row 9 can be written as:

$$R_T = \frac{P_T}{\frac{P_S}{R_S} + \frac{P_N}{R_N}} \qquad [1]$$

where

S = subscript for SSI recipient
N = subscript for non-SSI recipient
T = subscript for total across above categories
R = participation rate (Rows 1, 2 and 9)
P = number of participants (Rows 3-5)

The participation estimates P_T, P_S, and P_N are taken from program data and, as an approximation, are assumed not to have sampling variance.
It can be shown as a theorem in statistics, that if K is a constant and X is a random variable,

$$\mathrm{Var}\left[\frac{K}{X}\right] \sim \left[\frac{\bar{K}}{\bar{X}}\right]^2 \left[\frac{\mathrm{Var}\,(X)}{\bar{X}^2}\right] \qquad [2]$$

where a variable with a line over it represents a mean. (See Mood, et al., 1974.) Application of this to equation (1) yields

$$\mathrm{Var}\,[R_T] = \frac{P_T^2}{\left(\frac{P_S}{R_S} + \frac{P_N}{R_N}\right)^2} \cdot \frac{\mathrm{Var}\left(\frac{P_S}{R_S} + \frac{P_N}{R_N}\right)}{\left(\frac{P_S}{R_S} + \frac{P_N}{R_N}\right)^2} \qquad [3]$$

Table of Contents

Applying equation [2] again yields

$$\text{Var}\,[R_T] = \frac{P_T^2}{\left(\frac{P_S}{R_S} + \frac{P_N}{R_N}\right)^2} \cdot \frac{\left(\frac{P_S}{R_S}\right)^2 \frac{\text{Var } R_S}{R_S^2} + \left(\frac{P_N}{R_N}\right)^2 \frac{\text{Var } R_N}{R_N^2}}{\left(\frac{P_S}{R_S} + \frac{P_N}{R_N}\right)^2} \qquad [4]$$

This equation, together with the standard errors of R_S and R_N shown in parentheses in Rows 1 and 2 of the table, was used to calculate the standard errors in parentheses in Row 9.

APPENDIX J: DETAILS CONCERNING DIETARY INTAKE DATA AND ANALYSIS

This appendix presents details concerning the dietary intake data and analysis.

Measured Dietary Intake Levels Compared with Levels Obtained in Other Surveys

Table J.1 presents average levels of nutrient intake for low income elderly persons, as measured by the survey done for the current project and by two other surveys: the Health and Nutrition Examination Survey (HANES) done in 1971-1974 by the U.S. Department of Health, Education, and Welfare; and the 1977-1978 Nationwide Food Consumption Survey done by the U.S. Department of Agriculture (USDA) (U.S. Department of Agriculture, 1982).

In general, the nutrient intakes measured in the current study are lower than those found in the other surveys. There are several possible reasons for this. First, it should be noted that the nutrient levels observed in the current survey are much closer to those obtained in the HANES survey than they are to the USDA totals. The average difference between the current survey and the HANES totals is only 5 percent as compared with 17 percent for the USDA survey. As discussed in Volume III, the interviewing protocols and data processing software used in the current survey were, for the most part, patterned after those used by HANES. Thus, it is likely that a substantial share of the differences between the results of the current survey and those of the USDA survey are not due to factors unique to the current survey, such as the use of a telephone interviewing methodology, but rather are due to differences between the HANES and USDA methodologies. It is not currently possible to determine whether the HANES or the USDA procedures are the more accurate.[1/]

Possible seasonality in consumption may also account for differences in observed nutrient intake in the current survey as compared with those of other surveys. Most of the interviews conducted for the current study were done during the summer of 1981, interviewing for the USDA survey was conducted during November 1977 to March 1978, and the HANES survey was conducted over several years. Interviewers reported that many respondents in the current study remarked that

---

[1/] It should be noted that the HANES data for most nutrients other than calories and protein may themselves underestimate current consumption levels. The reason is that the HANES data were collected in the early 1970s, and there is evidence from periodic Department of Agriculture surveys that consumption levels of most nutrients other than calories and protein have been rising over time. However, the HANES intake estimates are, in general, lower than those obtained in an earlier 1965-66 USDA survey done prior to HANES. This suggests that even after taking changing consumption patterns into account, there are differences between HANES and USDA procedures that lead to significantly different intake estimates.

Table of Contents

TABLE J.1

LEVELS OF DIETARY INTAKE FOR LOW INCOME ELDERLY PERSONS
AS MEASURED BY DIFFERENT SURVEYS

	Current Survey	1971-74 HANES [a] Survey	Preliminary Data From 1977-78 Dept. of Agriculture [b] Survey	% Difference Between Current Survey & HANES [d]	% Difference Between Current Survey & USDA [d]
WOMEN					
Calories (Kcal)	1178.75	1197.44	1288.59	-1.6	-8.5
Protein (gm)	45.26	49.02	55.14	-7.7	-17.9
Calcium (mg)	448.22	519.17	593.93	-13.7	-24.5
Iron (mg)	7.95	7.89	9.84	7.6	-19.2
Vitamin A (IU)	4615.92	4417.35	7589.40	4.5	-39.2
Thiamine (mg)	0.91	.88	1.00	3.4	-9.0
Riboflavin (mg)	1.17	1.24	1.32	-5.8	-11.4
Niacin (mg)	10.65	11.23	13.56	-5.2	-21.5
Vitamin C (mg)	77.29	71.82	70.27	7.8	10.0
Average difference				-1.2	-15.7
MEN					
Calories (Kcal)	1366.81	1672.07	1724.09	-18.2	-20.7
Protein (gm)	55.58	64.42	71.27	-13.7	-22.0
Calcium (mg)	516.11	597.48	649.22	-13.6	-20.5
Iron (mg)	8.82	11.25	12.49	-21.6	-29.4
Vitamin A (IU)	3896.90	4342.17	5310.19	-10.2	-26.6
Thiamine (mg)	1.00	1.16	1.23	-13.8	-18.7
Riboflavin (mg)	1.28	1.52	1.53	-15.8	-16.3
Niacin (mg)	12.11	14.05	16.73	-13.8	-27.6
Vitamin C (mg)	63.25	69.23	57.62	-8.6	9.8
Average difference				-14.4	-19.1
AVERAGE [c]					
Calories (Kcal)	1217.06	1294.26	1377.43	-6.0	-11.6
Protein (gm)	47.36	52.16	58.43	-9.2	-18.9
Calcium (mg)	462.05	535.14	605.21	-13.7	-23.6
Iron (mg)	8.13	8.58	10.38	-5.2	-21.7
Vitamin A (IU)	4469.47	4402.01	7124.44	1.5	-37.3
Thiamine (mg)	0.93	0.94	1.05	-1.1	-11.4
Riboflavin (mg)	1.19	1.30	1.36	-8.5	-12.5
Niacin (mg)	10.95	11.80	14.21	-7.2	-22.9
Vitamin C (mg)	74.41	71.29	67.69	4.4	9.9
Average difference				-5.0	-16.7

[a] U.S. Department of Health, Education, and Welfare (1979).

[b] U.S. Department of Agriculture (1982).

[c] Weighted averages, with weights based on proportions of men and women in the current survey data (20.4 percent men and 79.6 percent women).

[d] Percentages are computed using the government survey as the base.

it was "just too hot to eat" when asked about their food consumption,[1] and this could have had a downward effect on food consumption, particularly with regard to calorie and protein intake.[2] Evidence that this may have been the case is provided by Table J.2, which shows differences in food intake between interviews covering days where the high temperature was 85 degrees or more as compared with days when the high temperature was less than 85 degrees. As shown in the table, intake was lower on the high-temperature days for each of the nine nutrients, and for seven of the nine nutrients the differences are statistically significant. Overall, the average percentage difference between the higher-temperature days and other days was approximately 11 percent.

Forty-five percent of the interviews in the sample were conducted on days with temperatures above 85 degrees. Thus the data suggest that, on average, nutrient intake recorded in the survey may have been approximately 5 percent [.45 times 11 percent] lower than it would have been if none of the interviews had been given on days with high temperatures.

It should be noted that, strictly speaking, these data cannot be interpreted as directly showing the effect of having conducted the interviews over the summer. Rather these data show intra-day variation within the summer months. The tabulations thus demonstrate that within the summer months, hotter days tend to lower consumption, but they do not provide direct evidence regarding the possibility that overall patterns of nutrient intake may be lower [or higher] in the summer as compared with other times of the year. It is possible at the conceptual level that the effect of having interviewed during the summer could be either greater or lesser than the 5 percent estimate suggested by the above tabulations. Nevertheless, the data are at least consistent with the possibility that observed levels of intake were lower because of summer interviewing.

Another factor that should be noted is that the sample of elderly persons for the current study is somewhat different from the sample for which USDA survey data are available. The available USDA data include all elderly persons with low income, while the current survey was limited to elderly persons living in households with no members under 65 years old and who are eligible for food stamps. It seems likely that elderly persons may, on average, have access to more and better food when they are living in larger households which include younger members as well.

A final possibility, however, is that some food consumption may have been underreported in the current survey. There is no way to determine with certainty whether this is the case. It is important to note, however, that even if some underreporting did occur, it is likely that it did not affect

---

[1] This was particularly true at the Oregon site, which experienced record high temperatures during parts of the survey period.

[2] Partially offsetting negative effects of the heat could have been possible positive effects from the availability of fresh fruits and vegetables during the summer months.

Table of Contents

TABLE J.2

DIFFERENCES IN NUTRIENT INTAKE FOR DAYS WITH HIGH TEMPERATURES, 85 DEGREES OR MORE

	(1) Intake on Days with High Tempera-tures Below 85	(2) Intake on Days with High Tempera-tures 85 and over	(3) Differ-ence	(4) Difference as Percentage of Intake on Days below 85
Calories (Kcal)	1257.01	1168.49	-88.52* (2.72)	7.0%
Protein (gm)	50.11	44.03	-6.08* (4.09)	12.1
Calcium (mg)	499.81	416.38	-83.23* (4.62)	16.7
Iron (mg)	8.44	7.76	-0.68* (2.57)	8.1
Vitamin A (IU)	4818.19	3923.93	-894.26* (2.46)	20.2
Vitamin C (mg)	76.43	71.97	-4.46 (1.15)	5.8
Thiamin (mg)	0.94	0.91	-0.03 (1.25)	3.2
Riboflavin (mg)	1.27	1.10	-0.17* (3.05)	13.4
Niacin (mg)	11.53	10.24	-1.29* (3.38)	11.2
Average Percent Difference				10.9

NOTES: Entries are units of nutrient.

Absolute values of t statistics are shown in parentheses under entries in Column (3).

Asterisks indicate that estimated differences are statistically significant with a .05-level two-tailed test.

any of the key conclusions of the analysis. The focus of the analysis is on comparisons of dietary intake between groups of individuals, such as comparisons between program participants and nonparticipants or comparisons between participants receiving cash and participants receiving coupons. Even if some underreporting occurred in the survey, there is no reason to believe that it would have occurred differentially more among some of these groups rather than others.

Probability of Meeting RDA

As discussed in Chapter VIII of the report, 24-hour recall data based on a single day of food consumption do not provide accurate information with regard to proportions of the population meeting recommended daily allowances (RDAs) of nutrients. (See Chapter VIII for a discussion of the reason for this.) As a result, the analysis presented in the text of the report does not focus on RDAs as an outcome measure. However, because there may be some interest in the RDAs observed in the survey, this section presents tabulations of percentages of respondents who met RDAs and also presents the results of probit analysis of RDA outcomes.

The RDAs used for the analysis are those developed by the National Academy of Sciences (1980).[1/] As background for the analysis of probabilities of meeting RDAs, it may be useful to examine the relationship between average intakes as measured by the current survey and the RDAs. Table J.3 presents these data. In general, the average intakes observed in the sample are lower than the RDAs.

Tables J.4 through J.6 present data on proportions of households meeting RDAs for the comparison and demonstration site samples. As with the nutrient intake comparisons in Chapter VIII, the RDA results are presented for the raw nutrient data and also with the effects of other variables controlled using probit. The independent variables used in the probit equations are similar to those used in the regression equations reported in Chapter VIII. (Complete probit results are included in Appendix M.) The probit results reported in the fourth column of each table can be interpreted as the percentage difference in the likelihood of a respondent meeting the RDA for a given nutrient after controlling for other variables. For example, in the raw data, participants in comparison sites had a .035 lower likelihood of reaching the calorie RDA than nonparticipants (Table J.4, Column 3). The difference in probabilities changes to an estimated .040 lower probability when variables other than participation are controlled for.

The patterns of results are generally similar to those found in the analysis of program effects on average nutrient intakes. At comparison sites, participation generally had a small and negative, but statistically insignificant effect on the likelihood of a respondent meeting nutrient adequacy standards. At cashout sites, participation had a generally positive effect on the likelihood of a respondent's diet meeting adequacy levels, and for five of the nutrients, the

---

[1/] Except for calories, RDA levels are set in such a way that meeting the RDA for a nutrient will provide sufficient intake for 95 percent of the population. (The comparable percentage for calories is 50 percent.) Thus, failure of an individual to meet an RDA level does not necessarily mean that the person is consuming an inadequate amount of the nutrient, given that persons' own requirements.

Table of Contents

TABLE J.3

AVERAGE NUTRIENT INTAKE AS PERCENTAGE OF RDAs

	Current Survey Average Intake	RDA	Average Intake as Percentage of RDA
WOMEN			
Calories (Kcal)	1178.75	1800[a/]	.65
Protein (gm)	45.26	44	1.03
Calcium (mg)	448.22	800	.56
Iron (mg)	7.95	10	.80
Vitamin A (IU)	4615.92	4000	1.15
Thiamine (mg)	0.91	1.0	.91
Riboflavin (mg)	1.17	1.2	.98
Niacin (mg)	10.65	13	.82
Vitamin C (mg)	77.29	60	1.28
MEN			
Calories (Kcal)	1366.81	2400[a/]	.57
Protein (gm)	55.58	56	.99
Calcium (mg)	516.11	800	.65
Iron (mg)	8.82	10	.88
Vitamin A (IU)	4896.90	5000	.98
Thiamine (mg)	1.00	1.2	.83
Riboflavin (mg)	1.28	1.4	.91
Niacin (mg)	12.11	16	.76
Vitamin C (mg)	63.25	60	1.05

[a/] Calorie RDAs shown in the table are midpoints of ranges for persons 51-75 years old. In the probit analysis, for persons older than 75, the midpoints of the range for persons older than 75 were used. These are 1,600 and 2,050.

TABLE J.4

DIFFERENCES IN PROBABILITY OF MEETING RDAs
BETWEEN PARTICIPANTS AND NONPARTICIPANTS

COMPARISON SITES

	(1)	(2)	(3)	(4)
	Raw Data			
	Participant	Non Participant	Difference	Difference After Controlling for Effects of Other Variables
Calories				
Protein				
Calcium				
Iron				
Vitamin A				
Vitamin C				
Thiamin				
Riboflavin				
Niacin				

NOTES: Entries are probabilities.

Absolute values of t statistics are shown in parentheses under entries in Column (4).

See Appendix M for complete probit results.

TABLE J.5

DIFFERENCES IN PROBABILITY OF MEETING RDAs
BETWEEN PARTICIPANTS AND NONPARTICIPANTS

CASHOUT SITES

	[1]	[2]	[3]	[4]
	Raw Data			
	Participant	Non Participant	Difference	Difference After Controlling for Effects of Other Variables
Calories				
Protein				
Calcium				
Iron				
Vitamin A				
Vitamin C				
Thiamin				
Riboflavin				
Niacin				

NOTES: Entries are probabilities.

Absolute values of t statistics are shown in parentheses under entries in Column [4].

Asterisks indicate that estimated effects are statistically significant with a .05 level two-tailed test.

See Appendix M for complete probit results.

Table of Contents

TABLE J.6

DIFFERENCES IN PROBABILITY OF MEETING RDAs
BETWEEN PARTICIPANTS AND NONPARTICIPANTS

TOTAL DATA SET

	(1)	(2)	(3)	(4)
	Raw Data			Difference After Controlling for Effects of Other Variables
	Participant	Non Participant	Difference	
Calories				
Protein				
Calcium				
Iron				
Vitamin A				
Vitamin C				
Thiamin				
Riboflavin				
Niacin				

NOTES: Entries are probabilities.

Absolute values of t statistics are shown in parentheses under entries in Column (4).

Asterisks indicate that estimated effects are statistically significant with a .05 level two-tailed test.

See Appendix M for complete probit results.

estimated effects are statistically significant. The differences in the effects of participation between comparison and cashout sites may be due to sampling error since there is no inherent reason to believe that cash, which reduces the link between participation and food or nutrient intake, is likely to improve dietary adequacy. When data are pooled across sites, most of the estimated effects are positive, but only one—that for protein—is statistically significant.

The Self Selection Correction Factor

As discussed in Chapter VIII, tests were undertaken in which a correction factor based on an estimated probit model of participation was used to control for possible self-selection bias in the nutrient intake regressions. The procedures used were based on Heckman (1979). Since inclusion of this factor did not substantially alter the results of the analysis, the factor was not included in the final equation specifications on which the results reported in the main body of the report are based. This section describes in more detail the work in this area and summarizes results of equations estimated with the correction factor included in the specification.

Let d be a 1,0 variable indicating whether an observation is a Food Stamp Program participant. Then a probit model of the participation process can be specified as,

$$d = 1 \text{ if } \underline{X}' B + v > 0,$$

$$d = 0 \text{ if } \underline{X}' B + v < 0,$$

where the X variables are determinants of participation, B is a vector of estimated parameters, and v is an error term, with an estimated standard error S_v.

The correction factor, c, inserted into the nutrient intake regressions was computed as

$$c = \frac{\text{normal density function of } [X'B/S_v]}{\text{cumulative normal density function of } [[2d-1][X'B/S_v]]}$$

Table J.7 summarizes the effects of including this correction factor in the nutrient intake regressions. The results are based on regressions for the combined sample pooled across all comparison and demonstration survey sites. The first column in the table shows estimated coefficients on the 1,0 indicator of program participation, for equations without the correction factor.[1/] The second column of the table presents comparable results for equations estimated with the correction factor. Absolute values of the t statistics associated with the coefficients are shown in parentheses. As shown in the table, the results are not substantially altered by the inclusion of the correction factor. Most of the estimated coefficients are very small in

---

[1/] These numbers are from Table VIII.8 of the main body of the report.

TABLE J.7

EFFECTS OF INCLUDING SELF-SELECTION CORRECTION FACTOR IN NUTRIENT INTAKE REGRESSIONS
[Data for all sites pooled]

	Coefficient on Program Participation in Equation Without Correction Factor	Coefficient on Program Participation In Equation With Correction Factor
Calories [Kcal]		
Protein [gm]		
Calcium [mg]		
Iron [mg]		
Vitamin A [IU]		
Vitamin C [mg]		
Thiamin [mg]		
Riboflavin [mg]		
Niacin [mg]		
Sample Sizes in regressions[a/]		

NOTE: Absolute values of t statistics are shown in parentheses.

a/ Sample size is lower in regressions with correction factor because some observations lacked the data needed to compute this variable.

relation to their standard errors and remain so when the correction factor is included. To be sure, the absolute values of some of the coefficients change substantially. For instance, the estimated coefficient in the calories equation changes from 3.72 to 23.3. However, this generally happens in instances where both estimated coefficients are quite small relative to the average values of the dependent variables and relative to their standard errors. In the case of calories, for instance, even the larger estimate is less than 3 percent of average caloric intake in the sample and is much smaller than its standard error.

Table of Contents

APPENDIX K: REPORTED INCOME: SURVEY COMPARED WITH CASE RECORDS DATA

Income data reported by Food Stamp Program participants in the survey were compared with case records data for the same households to provide some indication of the degree of underreporting that occurred. To be sure, it must be recognized that the case records themselves are likely to be subject to considerable error. Therefore, not all discrepancies between the two data sets should be attributed to errors in the survey information. Nevertheless, comparison of the two types of data can at least be indicative of whether the survey data are similar to those that would have been obtained by eligibility workers during actual program application recertification interviews.

The names of sample members who were found during the survey to be program participants were matched against case records listings supplied by the sites. In cases where apparent matches were identified, the case records data were combined with the survey data on a single analysis file. Cases where there appeared to be substantial discrepancies in household demographic data such as race, age, or sex of the head of the household were eliminated from the file for the matched analysis, on the grounds that such discrepancies may have been indicative of incorrect matching. The analysis used data from both New York sites and from the demonstration sites in South Carolina and Oregon.[1/]

The overall sample size available for the analysis involving Food Stamp Program benefit levels was 651 cases. Somewhat fewer cases were available for comparisons of gross and net income levels, because some sites did not include these data in the case records information.

One limitation with regard to the comparison of survey and case records income data should be noted: federal SSI and Social Security benefit levels were increased by 11.2 percent as of July 1, 1981. The survey began at approximately the same time, and thus the survey data reflect the increases in federal benefit levels for these programs as of that date. However, the case records data for this analysis were supplied in the late summer and early fall of 1981. This means that the most recent Food Stamp Program recertification for many of the households in the data set had occurred prior to July 1. For such households, the case records data do not reflect the July 1 SSI and Social Security increases and therefore underestimate income receipts as of the time of the survey.

The available case records data do not allow a determination of the precise magnitude of income undercounting in the case records due to this factor. However, the amount of undercounting is certainly considerably less than the 11.2 percent federal benefit increase. There are several reasons for this: (1) rises in state SSI benefits were lower than the increase in federal benefits;

[1/] Because of a programming error, case records data for the appropriate time period were not available for the South Carolina and Oregon comparison sites.

(2) some of the respondent households have other income sources besides SSI and Social Security; and (3) some of the cases had been recertified after July 1 and, for such households, the case records data reflect the July 1 increase. Thus, while there is some undercounting in the case records data due to the timing of the increase in federal benefits, the extent of the undercounting can be assumed to be under 11.2 percent.

Table K.1 summarizes results of the analysis. As shown in the table, there is considerable variation between income estimates in the survey data and in the case records data. Only 33 percent of the gross income estimates and 21 percent of the net income estimates are within $10 of each other in the two data sets.[1/]

As indicated in the bottom row of Table K.1, reporting discrepancies tend to offset one another, so that when averaged over all of the cases on the file, they are relatively small. The average discrepancies for both gross and net income are under $4 and are smaller than their standard errors. The average net discrepancy for food stamp bonus amount is $1.6, and this difference is statistically significant.[2/]

As shown in Table K.2, there is considerable variation by site in the size and nature of the discrepancies in the data. The two New York State sites have the lowest average errors. Average net income is $4.2 higher in the survey data than in the case records data for the New York demonstration site and $21 higher at the comparison site. The discrepancies in net benefit amounts at these sites are about $4 at the demonstration site and -$2 at the comparison site.

Average discrepancies are substantially larger at the South Carolina and Oregon sites, but the differences are largely offsetting. Average net income as reported in the survey data is $36 lower than the case records data, and average benefits are $13 higher at the South Carolina site. At the Oregon site, on the other hand, average survey income is $43 higher, and average benefits are $12 lower than the corresponding case records information.

[1/] At first examination, it may appear surprising that the discrepancies are larger for net income after deductions than they are for gross income. However, two factors may at least in part account for this. First, net income is computed as gross income minus deductions. Thus, in computing net income there is both the possibility of potential error in estimating gross income and an additional source of potential error stemming from discrepancies in deductions estimates. Second, because of the way in which deductions are calculated in computing net income for the program, errors in gross income tend to be compounded when estimating net income. The reason is that the housing deduction is computed as actual housing costs in excess of half of income after other deductions have been subtracted. Thus, if an error is made in measuring gross income, it can lead to the opposite error in estimating the housing deductions. The error is then compounded when deductions are subtracted from gross income.

[2/] The survey-based estimates of food stamp bonus amounts used in the analysis were calculated from survey data on income and deductions. Discrepancies in bonus amounts are therefore correlated with income discrepancies.

Table of Contents

TABLE K.1

DISCREPANCIES BETWEEN SURVEY AND CASE RECORDS DATA

Size and Direction of Discrepancy	Gross Income[a]	Net Income After Deductions[b]	Food Stamp Benefit Amount[c]
Survey higher by >$50	4%	20%	4%
Survey higher by 41–50	4	4	2
Survey higher by 31–40	4	6	3
Survey higher by 21–30	9	5	6
Survey higher by 11–20	15	6	8
Discrepancy $10 or less	33	21	55
Survey lower by $11–20	9	6	10
Survey lower by 21–30	6	4	6
Survey lower by 31–40	3	4	3
Survey lower by 41–50	2	3	2
Survey lower by >$50	11	19	2
AVERAGE DISCREPANCY	-$2.2[d]	$3.1	$1.6
	(3.4)	(3.7)	(.8)

[a] Based on sample of 394 matched records. Data were not available for the two New York sites.

[b] Based on sample of 564 matched records. Data were not available for some cases at the Monroe County, New York site.

[c] Based on sample of 650 matched records.

[d] Standard errors of average discrepancies are shown in parentheses.

Table of Contents

TABLE K.2

DISCREPANCIES BETWEEN SURVEY AND CASE RECORDS DATA, BY SITE

	Sample Size	Gross Income (Survey–Case Records)	Average Discrepancy in Net Income After Deductions (Survey–Case Records)	Average Discrepancy in Food Stamp Benefit Amount (Survey–Case Records)
New York Demonstration Site	155	NA	$4.2[a]	$3.7
			(11.3)	(1.6)
New York Comparison Site	101	NA	21.0	-2.2
			(8.0)	(1.8)
South Carolina Demonstration Site	222	$-11.8	-36.3	12.5
		(5.8)	(5.9)	(1.3)
Oregon Demonstration Site	172	10.2	43.1	-12.2
		(2.3)	(5.1)	(1.5)

NOTES: Standard errors of estimated averages are shown in parentheses.

NA = not available.

[a] Net income data were available for only 69 observations at the New York demonstration site.

Overall, the results of the matched survey data/case records analysis indicate considerable discrepancies between the survey data and the case records data. While neither information source can be assumed to be completely correct, it is reasonable to believe that information collected during program certification interviews, where there are legal requirements to provide accurate data, is probably more accurate.

To a large degree, the discrepancies between the survey and case records data tend to be offsetting, on average, and average income and benefit levels are quite similar between the survey and the case records information. However, as noted earlier, the income data in the case records are themselves underestimates of true income at the time of the survey because the case records do not fully reflect the July 1, 1981 increases in SSI and Social Security payments. In light of this, the fact that the two data sources provide similar average estimates suggests that there is, on average, some underreporting in the survey data. However, the amount of the underreporting is probably under 11 percent. It therefore seems unlikely that any of the major conclusions of the analysis have been substantially affected by errors in the survey data.

Table of Contents

APPENDIX L: PROGRAM ELIGIBILITY CALCULATIONS BASED ON RETROSPECTIVE INCOME DATA

This appendix presents technical details concerning the program eligibility calculations based on retrospective data used in support of the analysis presented in Chapter X of the report. Differences between the microsimulation work performed for the current project and a typical full scale simulation model are discussed, then details of how income and assets were simulated using Current Population Survey (CPS) data are given. Next, the program eligibility rates estimated from the current model are compared with an independent estimate of this rate. Standard errors for the net discrepancy rate estimates presented in Chapter X are then calculated, and the appendix concludes by summarizing reasons for sample attrition in the simulation analysis.

Current Approach Compared with Full Simulation Model

It should be emphasized that the procedures and the analysis presented in Chapter X and in this appendix cannot be considered a validation of currently-used simulation models because a number of sets of assumptions frequently used in models could not be tested. In addition, certain aspects of the CPS data base could not be replicated with the retrospective data obtained in the survey for the current project. The following frequently used simulation model assumptions were not examined:

(1) Data used in the simulation models are often several years old, and the models employ complex "aging" processes to project household and other data to a current basis. Because the current survey contained relatively few observations and was not nationally representative, the standard aging procedures were not used and thus could not be tested.

(2) Similarly, underreporting of income is often corrected for by adjusting income totals to known national totals. Because such control totals were not available for local survey sites, this aspect of simulation modeling could not be tested.

(3) In estimating allowable Food Stamp Program deductions, simulation models sometimes use sets of expense imputation procedures based on national data. The imputation equations may produce biased results when applied only to a subset of the population, so reported expenses were used to estimate allowable deductions for the survey data.

(4) Some simulation models simulate all major welfare programs, as well as the Food Stamp Program, and these other simulated payments are used as input to food stamp eligibility determinations. This was not tested in the current work.

(5) Finally, because the CPS is a publicly distributed file, considerable resources are allocated by the Census Bureau toward resolving data inconsistencies and correcting for survey nonresponse. The current analysis was restricted to only those observations in which data appeared to be properly reported.

ELIGIBILITY SIMULATION

The microsimulation technique applies to individual micro units or observations from a survey (households in this case) a set of program rules that simulates eligibility and benefits for each unit—much in the same way a caseworker would determine the eligibility of a given applicant. Although the computations are performed on the individual units, microsimulation results are only used in the aggregate. That is, summary statistics of the total numbers of eligibles and participants are prepared, from which program participation characteristics are examined. The assumption is that the simulation results are accurate, on average, thereby producing reliable summary results. However, it can easily be demonstrated that for specific observations, the results are often incorrect. Thus, the objective is to determine the overall accuracy of the eligibility determination process rather than that of any individual household.

To afford a careful comparison, the methodology that was used in the current study is presented alongside procedures that might typically be used in a larger model to simulate eligibility using retrospective income data. The detailed algorithms developed for the present study are then given.

The following steps, for example, might be taken to simulate eligibility on the March 1981 CPS:

(1) Obtain a data file from Census for which consistency edits and imputations for nonresponse have already been performed.

(2) Allocate income reported in combined source categories to the individual components.

(3) Alter the retrospective labor force data to be consistent with the survey week data.

(4) Age the income to reflect calendar year 1981.

(5) Possibly correct certain income types for survey underreporting and nonreporting.

(6) Simulate public assistance and SSI.

(7) Calculate monthly income.

(8) Simulate eligibility under the Food Stamp Program.

The procedures followed with the current data set and deviations from the standard procedures are discussed below.

Data Editing. The survey data file contained a number of observations for which at least one component of annual income could not be determined accurately. The current sample size was too small to permit reliable estimates of mean values for missing data; therefore, imputations were not made. All cases where annual income amounts could not be determined from the reported data were screened out of the analysis.

Income Allocation. As in the CPS, annual income items on the survey were collected by first asking if an individual had received certain types of income and, if so, the amount. For some items the recipiency question referred to a single source, whereas for other items, there was a list of two, three, or four sources; the amount reported was the sum received from all of them. In a full microsimulation, these sum amounts are routinely allocated to component sources, because amounts received from various sources are treated differently. For the simulation of food stamp eligibility, however, the only source treated differently from the other components with which it was combined was unemployment compensation. In fact, this source was seldom reported in the survey. The only case where it was reported jointly with other components was eliminated due to nonresponse concerning the amount received. Hence, no allocation of income was performed for this analysis.

Labor Force Data Adjustment. The CPS collects data on labor force activity during the survey week [the second week in March] as well as activity during the previous calendar year. There is the potential for inconsistency between these data items because people who worked during the entire year may have left the labor force prior to March of the subsequent year, or the reverse could happen. In a full microsimulation model, this potential conflict can be resolved for a typical food stamp eligibility simulation by using a labor force adjustment algorithm. Such an algorithm could not be applied in the current context because the full battery of CPS labor force activity questions was not duplicated in the current survey. Also, the adjustment process requires independent data on unemployment and labor force participation rates, which could not be obtained separately for the survey sites. In light of these factors, as well as the low labor force participation rate among the survey population, this step was omitted.

Aging the Data Base. In order to perform comparisons across the different program simulations, a consistent set of program parameters, and hence, a consistent time frame, must be used. So that the current monthly income concept would require minimum data manipulation, the time period chosen for this simulation was July of 1981. Were this analysis to be performed with the March 1981 CPS, the data base would ordinarily be aged so that it reflected the economic and demographic conditions in effect during the twelve months surrounding July [i.e., calendar year 1981]. This would involve adjusting the labor force data as mentioned in the preceding step and aging the income data reported at the person level to reflect the income levels the sample population was expected to receive in 1981.[1/]

[1/] Typical aging procedures also include altering the demographics to reflect those of the population existing at the middle month of the simulation year [calendar year 1981 in this case]. However, that step would not be needed in the present case, because March is sufficiently close to the middle month of the year.

As discussed above, the labor force data were not adjusted for this project. However, each reported income amount was aged by inflating it by a factor that represented the expected change in the level of income receipt over the period in question. The aging factors were derived separately for each source. To the extent possible, they were based on observed changes in average income received by the elderly population over the period 1980 to 1981. The individual aging factors used are described below.

ADJUSTMENTS

In general, the March CPS files, after editing and imputation, have been found to underreport income received by the household sector [Doyle, et al. 1980]. Therefore, some microsimulation systems have optional procedures that adjust the amounts reported in individual records so that in the aggregate, total income from the component sources equals amounts estimated from independent sources. Because independent control data for the current survey sites could not be obtained, this step was omitted in the present work.

Simulation of Public Assistance and SSI. Simulation models are often designed to produce estimates of receipts from the major means-tested cash transfer programs in addition to food stamps. Therefore, the public assistance and SSI data used as input to the food stamp eligibility determination are sometimes the result of a fairly complex microsimulation model that simulates the participation decision and applies program rules to determine eligibility and benefits. Because the focus of the current study was specifically on Food Stamp Program eligibility, it was decided to determine benefits from the previous year's means-tested transfer income rather than from simulated results.

However, the decision not to undertake public assistance simulation did not eliminate this step. It was still important to attempt to measure intra-year income streams, because doing so can have a significant effect on food stamp eligibility determination. Furthermore, the measurement error associated with using an approximation of intra-year income flows represents the type of measurement error generated by the use of annual retrospective income, which is one of the issues studied here. A complete description of the procedures used to simulate public assistance and SSI is presented below.

Calculation of Monthly Income. Neither the CPS itself nor the CPS portion of the current survey, contain much information on intra-year income flows. Error is thus introduced for observations with high turnover in the labor market and for those with irregular receipt of unearned income. In order to overcome these data limitations, microsimulation methods sometimes use both current labor force data and the retrospective annual income reported for each person to construct monthly income amounts at the individual level. Household monthly income is then the sum of these amounts across individuals within a household.

With regard to this analysis, it was determined whether each person in the sample was working during the simulation month, and earned income, SSI, and public assistance benefits were allocated accordingly. Other unearned income was allocated evenly throughout the year. The algorithms used to construct monthly amounts for these income sources are described below.

Simulation of Food Stamp Eligibility. Three calculations of eligibility were required for the analysis: one using prospective monthly income and reported assets; one using simulated monthly income and reported assets; and one using

simulated income and a proxy for assets. Except for the procedures noted below, the methods used to determine eligibility and benefits with the retrospective data were the same as those used in the determination of eligibility carried out for the other analyses documented in this report. Only the exceptions are described here. The basic eligibility algorithm is described in Appendix F.

In the procedure used for simulating eligibility from simulated monthly income and reported assets, there were two deviations from procedures used with current prospective income. The first exception was that the income used was the result of the monthly income calculation based on retrospective annual income. The second exception was that the household, rather than the food stamp unit, was the unit of analysis. This decision was made because the food stamp unit is not known with the CPS. However, the two concepts differed for only one household in the final sample.

The procedure used for the eligibility determination based on simulated income with the assets proxy was the same as in the preceding method with the exception of the assets test. The level of assets was calculated as the sum of income from interest, dividends, rents, royalties, and estates and trusts, divided by an average rate of return on investment. This computed level of assets was then compared with the program limits for elderly households in effect for July 1981, which were $1500 for a one-person unit and $3000 for a unit containing two or more persons. The rate of return on investment was set to 5.25 percent, which was the rate of return on passbook savings at the time. This rate of return was used because it was believed that most of the survey respondents primarily possessed only small amounts of savings that they tended to keep in passbook savings accounts.

DATA BASE PREPARATION

As described above, the procedure for preparing the data for the analysis reported in Chapter X consisted of the following steps:

(1) Age retrospective annual income to reflect calendar year 1981 dollars.

(2) Allocate public assistance and SSI income to periods of work and nonwork during the year.

(3) Calculate monthly income.

(4) Simulate eligibility for the Food Stamp Program.

The first three steps are discussed in detail below. The procedure used to simulate food stamp eligibility was similar to the procedure used with prospective income presented in Appendix F.

Aging Retrospective Income

Income was aged to calendar year 1981 by applying growth rates that vary by income source. To the extent possible, the growth rates were derived from data relevant to the elderly population. The rates applied to the individual income amounts and their sources are described below.

Earnings. Earnings, which represent the sum of wages and salaries, and farm and nonfarm self-employment income, were inflated by 10.2 percent. This is the

Intermediate II-B estimate for the increase in earnings in 1981 given in the July 1981 Social Security Trustee's Report (U.S. Senate Committee on Finance, 1981, Table 28).

Social Security. The second category includes benefits received from Social Security as well as those received from the Railroad Retirement Board. These benefits are indexed to the Consumer Price Index (CPI), with benefit increases effective July 1. Because the simulation came after the July 1981 increase, an inflation factor of 11.2 percent was used, representing the benefit increase effective that month (U.S. Senate Committee on Finance, 1981, Table 28).

Supplemental Security Income. For the three states in which the survey was conducted, eligible persons could receive up to a federal maximum SSI plus some state supplementation. For this analysis, increases of 8.9 percent for New York, 10.7 percent for Oregon, and 11.2 percent for South Carolina were assumed. These represent the statutory increases in the combined federal and state guarantees effective July 1981 (Social Security Administration, October 1980, and December 1981).

Public Assistance. Public assistance includes the Aid to Families with Dependent Children (AFDC) program as well as local general and emergency assistance programs. These are for the most part locally administered, and entitlements and the rates at which the guarantees changed over the study period varied significantly across the states surveyed. Therefore, three different multiplicative factors were applied: 5.7 percent for New York, 11.7 percent for Oregon, zero for South Carolina, based on discussions with state officials.

Interest. It is believed that the population surveyed mainly kept its savings in passbook savings accounts. Therefore, the aging factor used for interest represents the expected growth in interest income from calendar year 1980 to calendar year 1981. The maximum allowable interest rate for both calendar years 1980 and 1981 was 5.25 percent compounded quarterly. With that rate, the expected increase in interest income, assuming no deposits or withdrawals, is 5.35 percent.[1/]

Dividends. Dividends represent income received from several sources: dividends, net rents, royalties, and estates and trusts. If the present study were a project based on a large nationally representative survey, the aging factors would have been derived from macroeconomic data. However, it was felt that the survey population may not have experienced the same increase over time in amounts of this type of income as the general population. In the absence of data with which to estimate a more relevant inflation factor, this income was assumed to have increased at the same rate as interest income. Therefore an inflation factor of 5.35 percent was used.

[1/] The maximum allowable interest rate was determined from discussions with officials at the American Security Bank and Riggs National Bank, both in Washington, D.C.

Pensions. Income from private and government pensions, the latter including federal civil service, military retirement, and state and local pensions, has been observed to increase more slowly, on average, than the cost of living. For purposes of the current analysis it was assumed that the increase in pensions was equal to one-third of the change in the CPI from 1980 to 1981.[1/] The CPI rose 11.1 percent over the period of interest [U.S. Senate Committee on Finance, 1981, Table 28], so an inflation factor of 3.7 percent was used.

Compensation and Other Income. The final category includes Veterans' Compensation, Workers' Compensation, Unemployment Compensation, alimony, regular contributions from sources outside the household, and miscellaneous money income. In the absence of detailed data on changes in the level of these income receipts over time, it was assumed that they changed in accordance with the CPI and therefore increases of 11.1 percent were assumed [U.S. Senate Committee on Finance, 1981, Table 28].

Public Assistance and SSI Payments

The simulation that was done divided the year into two parts when simulating public assistance and SSI, one period during which earnings were received by the unit [weeks worked period] and one period during which earnings were not received [weeks not worked period]. Because levels of income receipt varied significantly across these two periods, separate public assistance [PA] and SSI benefits were computed for each period. The total annual benefit from these programs is the sum of the two part-year benefits.

Some observations reported total annual benefits and others reported average monthly amounts and months of receipt. In the latter case, the annual amount was constructed and used; if either the amount or the period of receipt was missing, the case was eliminated. SSI recipients who reported annual benefits in excess of $4000 and PA recipients who reported annual benefits in excess of $350 also were screened out. In order to allocate the annual benefits to the two part-year components, the number of months in which a case was not working was first compared with the reported number of months in which benefits were received. If the months of receipt did not exceed the months of non-work, all reported benefits were assigned to the weeks not worked time period and zero benefits were assigned to the weeks worked time period. This assumed that for these means-tested transfer programs, countable income was less during the non-work period and hence, the probability of receiving the transfer was greater. For cases where weeks worked encompassed the full year, all benefits received were allocated to the weeks worked time period. In all the remaining cases there was some evidence that the individuals received benefits during both periods of work and non-work. Due to the existence of earnings during the weeks worked period and the assumptions regarding the flow of other unearned income sources [discussed below], it was assumed that the average monthly benefit for

---

1/ This decision was based on two studies: (1) Gayle B. Thompson, "Impact of Inflation on Private Pensions of Retirees, 1970-74: Findings from the Retirement History Study." Social Security Bulletin, November 1978 and (2) Bankers Trust Company, 1975 Study of Corporate Pension Plans, 1975.

the weeks worked period would be lower than for the remainder of the year. Without data to determine directly how much lower those benefits should have been, amounts for the two periods were simulated, imposing the constraint that the sum of the two remain equal to the reported annual amount. The algorithms used for this were:

[1] For SSI

$$SSIWKN = MIN\ [SSI,\ [MAX[0,\ GUARS - Y_u/12 + 20]]\ *\ MWKN] \quad \text{if } Y_u/12 > 20$$

$$= MIN\ [SSI, [MAX[0, GUARS]] + MWKN] \quad \text{if } Y_u/12 < 20$$

$$SSIWKW = SSI - SSIWKN$$

where

SSI = Reported annual SSI benefits

SSIWKN = Amount of SSI allocated to the weeks not worked period.

SSIWKW = Amount of SSI allocated to the weeks worked period.

GUARS = Array of maximum monthly SSI benefits by state [328 for New York, 277 for Oregon and 265 for South Carolina].[1/]

Y_u = Other unearned income. Y_u is the sum of Social Security, interest, dividends, compensation, pensions and miscellaneous income.

MWKN = Months not working.

[2] For Public Assistance

$$PAWKN = MIN\ [PUBA,\ [Max[0, GUARP - Y_u/12]]\ *\ MWKN]$$

$$PAWKW = PUBA - PAWKN$$

where Y_u and MWKN are defined as above and

PUBA = Reported annual public assistance.

PAWKN = Amount of PA allocated to the weeks not worked period.

PAWKW = Amount of PA allocated to the weeks worked period.

[1/] Social Security Administration, 1981.

GUARP = Array of needs standards under the AFDC program by state [260 for New York, 277 for Oregon and 102 for South Carolina].[1/]

Calculation of Monthly Income

The general procedures used to construct monthly income for each observation varied according to whether a person were working during the simulation month. For those individuals who were employed, average monthly earnings were constructed along with average monthly SSI and PA benefits during the weeks worked period. For individuals who were not employed, monthly earnings were set equal to zero and average monthly SSI and PA benefits were calculated from the amounts allocated to the weeks not worked period. Average monthly other unearned income was constructed in the same way for both employed and unemployed people. Detailed methods for determining the components of monthly income are described below.

Monthly Earnings. The monthly earnings variable for employed persons was calculated as annual earnings divided by weeks worked converted to a monthly amount. Earnings represents the sum of income received from wages and salaries and from farm and non-farm self-employment. Cases where earnings were claimed but in which none of the three income sources had nonzero amounts reported were omitted from the study. Similarly, nonresponse to the question about weeks worked caused a case to be eliminated. Finally, for cases deemed to be employed but which had not worked in the previous year, monthly earnings were imputed based on the sample average.

SSI and PA. To construct the monthly SSI and PA amounts for employed persons, amounts allocated to the weeks worked period were first examined. If any of these conditions held:

[1] the amount received during the weeks worked period as estimated above was positive;

[2] the person had not received assistance during the previous year;

[3] the months receiving benefits in the previous year were less than or equal to the number of months not working, in which case the weeks worked period benefit was assumed to be zero;

then monthly SSI and PA amounts were calculated as the total benefit received during the months working period divided by the months receiving assistance during that period. Months receiving assistance during the working period was computed as the minimum of the number of months worked and the difference between the months receiving assistance and the months not worked.

[1/] Based on conversations with state officials. There were no published data on AFDC need standards as of January 1981. It would have been preferable here to use guarantees for the general assistance programs. However, they were not obtainable.

For employed observations for which the SSI payments allocated to the working period was 0, a monthly payment was simulated. The following algorithm was used for this simulation:

MSSI = Max[0,GUARS-{max [0,.5*{MEARN-65-[20-MUNER]}] if MUNER ≤ 20
= Max[0,GUARS-{max [0,.5{MEARN-65}] + (MUNER-20)}] if MUNER > 20

MPA = Max[0,GUARP-{max{0,.67*{MEARN-30}+MUNER}}]

where

GUARS and GUARP were defined in the previous section and

MSSI = Monthly SSI benefit.

MPA = Monthly PA benefit.

MEARN = Monthly earned income defined above.

MUNER = Monthly unearned income excluding means tested transfers.

To construct the monthly SSI and PA amounts for persons deemed not employed, the amount of benefits allocated to the weeks not worked was used and then converted to a monthly amount.

Unearned Income. Monthly unearned income was set equal to the sum of annual amounts reported from Social Security and railroad retirement, interest, dividends, compensation, pensions and miscellaneous sources divided by twelve. As was true with the other income amounts, cases that failed to respond to any of the unearned income questions were deleted from the study.

Aggregation Across Persons. The computations described above for earnings, SSI and PA benefits, and unearned income produced monthly person amounts. A further step of aggregating over the members of each household was then employed because the unit of interest is the household group when simulating food stamp eligibility. When these aggregates were derived, a flag was constructed denoting whether any household member was a nonrespondent in any of the variables discussed above. If so, the entire household was eliminated from the study.

COMPARISON OF ELIGIBILITY ESTIMATES WITH INDEPENDENT ESTIMATES

As a rough check on the validity of the adjusted data set, the estimated eligibility rates in Table X.1 of the main report were compared with independent estimates. From the weighted totals in Column 4, 150 units are eligible (the sum of Rows 1 and 2) using PIRA. This number of eligible units is 30.7 percent of the overall population represented in the table. With the RIRA simulation, 142 units are eligible (sum of Rows 1 and 3) which represents 29.1 percent of the overall population.

These estimates seem reasonable in light of available data about national eligibility rates for the elderly and about the incidence of poverty among the aged in the survey sites as compared with the country as a whole. Bickel et al. (1981) produced tabulations from Wave II of the 1979 Income Survey Development Program (ISDP) Research Panel survey showing that there are

approximately 4.[illegible] million household units in the United States that contain at least one member age 60 or older and that receive food stamps or are eligible for food stamps. This was 18.9 percent of such households.[1/] However, there is evidence that the elderly at the survey sites were poorer, on average, than the elderly in the country as a whole. The average across the survey sites of the percentage of elderly persons receiving SSI is approximately 1.5 times the national average.[2/] Receipt of SSI can be taken as an indicator of poverty among the elderly, and thus the incidence of poverty (and therefore, of Food Stamp eligibility) among the elderly at the survey sites may be, on average, approximately one and one-half times that for the whole country. This, together with Bickel et al.'s 18.9 percent approximate national Food Stamp Program eligibility rate estimate for the elderly cited earlier, suggests that the average across the six survey sites of the food stamp eligibility rate may be on the order of 28 percent. The eligibility rates from the current survey are reasonably consistent with this independently-derived estimates.

STANDARD ERRORS FOR NET DISCREPANCY RATE ESTIMATES

The text of Chapter X estimates the net discrepancy rate due to the use of retrospective income rather than current income data in estimating program eligibility as approximately 3.5 percent of the overall population of elderly households. This section calculates the standard error associated with that .01 estimate.

The net number of discrepancies in the fourth column of Table X.2 in Chapter X, i.e., 5, can be written as the weighted sum of the net numbers of discrepancies for the SSR and adjusted MBR samples, .08(3) + .92 (6). Similarly, the total number of cases eligible using retrospective data in Column 4 (Rows 1 and 3) is the weighted average for the numbers of cases in the two samples (.08)(285) + (.92)(133). Thus, the net discrepancy rate estimated for Column 4 (which is the estimated population net discrepancy rate) can be written as:

$$NDR = \frac{(.08)(3) + (.92)(6)}{(.08)(285) + (.92)(133)} = .035. \qquad (1)$$

1/ This is not strictly an eligibility rate because the numerator contains non-eligible recipients. Furthermore, the data file used in producing this figure contained preliminary sample weights. However, the 18.9 percent can be viewed as indicative of the national eligibility rate. See Czajka (1981) for a discussion of the issue. The 18.9 percent estimate is, in all likelihood, a lower-bound estimate of the relevant proportion for the current data set because entirely-elderly households (the population in the current data set) are probably poorer, on average, than households containing some younger members.

2/ As of the mid-1970s, approximately 10 percent of the elderly in the United States received SSI. The corresponding rates for the survey sites were: New York demonstration site, 6 percent; New York comparison site, 13 percent; South Carolina demonstration site, 29 percent; South Carolina comparison site, 31 percent; Oregon demonstration site, 4 percent; and Oregon comparison site, 4 percent (U.S. Bureau of the Census, 1977). The average across these six sites is approximately 15 percent, or one and one-half times the national rate.

Table of Contents

Within the SSR sample, the net number of three discrepancies is the difference of the rates at which the two possible discrepancies occur, times the total sample size for the SSR sample, i.e.,

$$3 = \left[\frac{9}{307}\right]-\left[\frac{6}{307}\right]\ [307]. \qquad [2]$$

Similarly for the MBR, the net number of six discrepancies can be written as:

$$6 = \left[\frac{17}{503}\right]-\left[\frac{11}{503}\right]\ [503]. \qquad [3]$$

Substituting [2] and [3] into [1] yields

$$NDR = \frac{[.08][307]\ \left[\frac{9}{307}-\frac{6}{307}\right]+[.92][503]\ \left[\frac{17}{503}-\frac{11}{503}\right]}{[.08][285]+[.92][133]}. \qquad [4]$$

Finally, because of the adjustment for MBR cases screened out on the basis of the mail interviews, the error rates of 17/503 and 11/503 for MBR cases can be written as the weighted averages of the rates for the cases actually on the data file and the cases artificially added to the data file to correct for the screening (assumed to have zero error). Thus

$$\frac{17}{503} = \left[\frac{280}{503}\right]\left[\frac{17}{280}\right] + \left[\frac{223}{503}\right]\left[\frac{0}{223}\right] \qquad [5]$$

and

$$\frac{11}{503} = \left[\frac{280}{503}\right]\left[\frac{11}{280}\right] + \left[\frac{223}{503}\right]\left[\frac{0}{223}\right]. \qquad [6]$$

Substituting this into [4] yields

$$NDR = \frac{[.08][307]\ \left[\frac{9}{307}-\frac{6}{307}\right]+[.92][503]\left[\left[\frac{280}{503}\right]\left[\frac{17}{280}\right]-\left[\frac{280}{503}\right]\left[\frac{11}{280}\right]\right]}{[.08][285]+[.92][133]}. \qquad [7]$$

Four of the rates in Equation (7) are rates at which the two types of discrepancies are estimated to occur in each of the two samples. These rates are all estimated from the data and are hence subject to sampling error. The remaining parameters in Equation (7) are weighting factors used to obtain population estimates, and as an approximation, these weighting factors will be assumed to be known with certainty in the variance calculations.

Therefore,

$$\begin{aligned} \text{var(NDR)} = & \left(\frac{(.08)(307)}{(.08)(285) + (.92)(133)}\right)^2 \text{var}\left(\frac{9}{307}\right) \\ & + \left(\frac{(.08)(307)}{(.08)(285) + (.92)(133)}\right)^2 \text{var}\left(\frac{6}{307}\right) \\ & + \left(\frac{(.92)(503)\left(\frac{280}{503}\right)}{(.08)(285) + (.92)(133)}\right)^2 \text{var}\left(\frac{17}{280}\right) \\ & + \left(\frac{(.92)(503)\left(\frac{280}{503}\right)}{(.08)(285) + (.92)(133)}\right)^2 \text{var}\left(\frac{11}{280}\right). \end{aligned} \qquad (8)$$

(Covariances between the estimated error rates can be assumed to be low because each of the rates is low, and as an approximation these covariances will be ignored.) Each of the variances in the equation can be estimated as variances of binomial distributions using

$$\text{Var} = \frac{(p)(1-p)}{\text{sample size}}.$$

Performing these calculations leads to the result that Var (NDR) = .0023, which implies a standard error of .048.

SAMPLE ATTRITION IN THE CPS-BASED ANALYSIS

The determination of the final sample size on which the analysis reported in Chapter X is summarized can be found in Table L.1. Interviews containing CPS data were conducted for 992 households. However, item nonresponse for income and assets data reduced the final available data set to 564 cases.

Table of Contents

TABLE L.1

REASONS FOR SAMPLE ATTRITION IN THE ANALYSIS
BASED ON CPS QUESTIONS

Completed interviews containing CPS data	
Interviews with missing data in CPS module	
Households with missing current prospective income data but who properly reported CPS data	
Analysis sample size	