THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

THE IMPACT OF IMPUTATION PROCEDURES
ON DISTRIBUTIONAL CHARACTERISTICS
OF LOW INCOME POPULATION

No. 8710 32

Pat Doyle and Robert Dalrymple

August 1987

U.S. Department of Commerce BUREAU OF THE CENSUS

ACKNOWLEDGEMENTS

CONTENTS

ABSTRACT

The analysis of the impact of the imputation procedures on the distributional characteristics of the low income population is concentrated on the population receiving benefits under the Food Stamp Program as reflected in the Survey of Income and Program Participation (SIPP). This is a means-tested in kind transfer program for which benefits are calculated as an explicit function of selected characteristics including income, unit size, presence of earnings and presence of elderly or disabled members. The existence of this explicit relationship provides a point of comparison for the outcome of the imputation process.

The analysis first examines the impact of imputation on aggregate characteristics of households receiving benefits and then examines the impact on aggregate characteristics of the subset of the household covered by the benefits (i.e., the food stamp unit). Noting a small impact of the imputation process on these aggregate statistics, the discussion shifts to a closer examination of the population whose benefits or income are imputed. In so doing, we show that the imputation process does not preserve the known relationship between benefits and the determinants of benefit levels.

In addition to analyzing the impact of imputation of benefits or income, we demonstrate that the distributional characteristics of the food stamp population are different when the unit of analysis is changed from the household concept routinely employed by the Census Bureau to the food stamp unit. We

further describe some problems encountered in the formation of food stamp units with SIPP.

I. INTRODUCTION

The Food Stamp Program is one of seventeen nutrition programs administered by the Food and Nutrition Service of the U.S. Department of Agriculture (FNS). This program provides coupons to the low income population which can be used to purchase food. In order to be eligible to receive food stamp benefits certain requirements must be met. These requirements are imposed on a group of people, hereafter referred to as the food stamp unit, who purchase and prepare meals together and who reside together.[1] The unit, once defined, is eligible to receive benefits if assets, gross income, and income exclusive of deductible expenses fall within specified limits which vary by unit size and geographic location. Units containing elderly (age 60+) or disabled members are allowed extra income and assets and higher deductions than other units with the same size. Once eligible, units' benefits are determined as an explicit function of income less deductible expenses and the Thrifty Food Plan.

The Survey of Income and Program Participation (SIPP) is a multi-panel longitudinal survey which collects information on Food Stamp Program participants as a part of the core questions repeated at every interview.[2] Also included in the core are questions pertaining to income which is countable under the

---

[1] There are some exceptions to this unit definition which allow the formation of separate units within the group of recipients sharing meals. A summary of the regulations which govern the formation of food stamp units is included in Appendix C.

[2] For an overview of SIPP see Nelson, et al. (1985).

1

program. Assets and most deductible expenses are measured in topical modules administered once (in the case of expenses) or twice (in the case of assets) over the two and a half year duration of each panel in the survey.[1]

SIPP is an important tool for the analysis of the Food Stamp Program (FSP) for a number of reasons. The collection of the determinants of eligibility noted above is one. Another important feature is that income and program participation are monthly statistics coincident with the program accounting period. A third feature is that an attempt is made to collect information on the food stamp unit within the Census definition of a household. Finally, the survey is longitudinal hence providing an opportunity to investigate the dynamics of the program participants.

A. OBJECTIVES

Before proceeding to use SIPP to analyze the FSP, FNS is interested in analyzing the quality of the information collected. The first step in this analysis was the comparison of SIPP-based estimates to estimates of program participants derived from surveys of administrative data on the program case-load (Dalrymple and Carlson, 1986). The second step, which is the subject of this paper, is to analyze the impact of the Census Bureau's nonresponse adjustments on the distributional characteristics of the low income population and to examine the implications of basing these characteristics on the Census household. Specifically we are concerned here with the impact of imputation

_____ _____

[1]The initial SIPP panel was interviewed over a two year and 10 month period. Subsequent panels were interviewed over a two year and 7 month period.

2

on the SIPP-based estimates of the economic characteristics of program participants and on the creation of the food stamp unit within the Census household.

1. Imputation

The Census Bureau deals with nonresponse in two different ways. First, the sample is weighted to compensate for most noninterviews. That is, most observations which refuse to participate are deleted from the sample and other observations have their sample weights adjusted so that when aggregated, the weighted totals reflect the full universe from which the sample was drawn. The second method of compensating for nonresponse is to impute data for the missing information in each incomplete record. The observations who refuse to participate but who remain in the sample have all of their responses to the questionnaire imputed. Individuals who are successfully interviewed but who failed to respond to one or more items in the questionnaire also have the missing items imputed.

Little is actually known about the true values of the items which are missing. The Census Bureau's imputation procedures are designed to preserve the overall mean and variance of the reported information for the total population. In the case of food stamps and other means-tested transfers, there is an explicit relationship between benefit levels and the determinants of benefit levels (income, unit size and composition, and earnings receipt) which is not taken into account in the Census Bureau procedures. Heeringa and Lepkowski (1986) note that imputation for a missing item may distort the relationship between it and other items unless specific controls are imposed. One objective of this paper is to see if in fact the absence of a control for the relationship

3

between benefit levels and the determinants of benefits levels does distort the relationship in the case of the Food Stamp Program.

We do not expect that the results of this analysis will tell us conclusively that the Census Bureau's imputation procedures are either "right or wrong." Instead, we expect to determine whether they are reasonable in light of the program regulations that govern the level of benefits relative to income, unit size, unit composition and presence of earnings.

One assumption made both by the Census Bureau in designing their imputation procedures and by us in testing the reasonableness of those procedures is that the reported information is accurate and internally consistent. This is a necessary assumption for the Census Bureau since their imputation methodology assigns reported values for nonreported values when household characteristics match. This is necessary for us since our comparisons utilize the relationships observed in the data, i.e., we do not develop a model of the Food Stamp Program and determine how well the imputed data fit that model. This assumption concerning the accuracy and consistency of the reported data may in fact not be correct for all observations in SIPP. There is a study currently under way at the Census Bureau to evaluate the quality of the reported data relative to administrative data for a sample of SIPP cases in four states (Kasprzyk, 1986). (This is referred to as the SIPP record check study). This study is designed specifically to explore the error rates in reported information in ten transfer programs including food stamps. The results of that study, when complete, will indicate whether the assumption regarding the correctness of the reported data is in fact valid.

4

2. Food Stamp Unit

The interview unit in SIPP is the Census household, i.e., a group of people who usually reside together. This interview unit is also a commonly used analytic unit when presenting SIPP-based statistics, particularly statistics on the Food Stamp Program. The food stamp unit, commonly called a household, is not in fact the same as the Census household. It is a subset of the Census household restricted to persons who purchase and prepare meals together, with some exceptions to allow for the formation of separate units for selected individuals such as elderly couples.[1]

SIPP does not measure all of the information needed to formulate a food stamp unit according to the program regulations. However, the survey does attempt to distinguish the food stamp unit within the Census household by soliciting information on which Census household members are covered by the benefit. As a result, we have observed that 18 percent of the Census households with food stamp benefits in September, 1983, contain at least one individual who is not covered by the benefit (Dalrymple and Carlson, 1986), and 16 percent of the Census households with food stamp benefits in both April and August 1984 contain a noncovered person.[2] The phenomenon of the existence of noncovered persons in food stamp households is explored more fully in Landa (1987).

_____

[1]The differences between the Census household and the food stamp unit are explained further in Appendix C.

[2]Based on tabulations of Waves 3 and 4 of the 1984 Panel of SIPP.

One objective of the paper is to examine the impact of the use of the food stamp unit instead of the Census household as the unit of analysis for the presentation of distributional characteristics of the food stamp population. It is important to note that while we can ascertain when income was imputed and when food stamp benefits were imputed, we cannot ascertain when unit composition was imputed. This is somewhat restricting because unit size is an important determinant of benefits given a specific income level and we simply have to assume that when benefits were reported, the unit composition was reported as well.

3. Glossary of Terms

The two concepts described above--imputation and unit of analysis--are used throughout the paper to define the universe in question. To assist the reader, the various universe definitions are defined below.

FOOD STAMP HOUSEHOLD	Census household containing at least one person who receives food stamps. Food stamp benefits may be reported or imputed in this case.
FOOD STAMP UNIT	Subset of the food stamp household consisting of only those people covered by food stamps. When more than one person in the Census household reported (or was imputed) food stamp benefits the unit consists of all persons covered by any of the benefits.
IMPUTED INCOME	The food stamp household or the food stamp unit (depending on the context) contains at least one person who did not report at least one income amount and the missing amount was imputed by the Census Bureau. The nonrespondent(s) could have refused to participate altogether (noninterview) or could have indicated receipt of a particular income type without supplying the

6

amount. Income refers to cash either in the form of earnings or unearned income.

IMPUTED BENEFITS

At least one member of the food stamp household or food stamp unit (depending on the context) indicated receipt of food stamp benefits but did not provide the face value of the coupons, which was imputed by the Census Bureau. This includes nonrespondent(s) who refused to participate altogether (noninterview) in which case receipt, benefit amounts and unit composition were all imputed. This also includes persons who reported receipt of food stamps but not their value. Unit composition may or may not have been imputed for such persons.

REPORTED INCOME

Households or units (depending on the context) who are not classified as having imputed income. Unit composition may or may not have been imputed.

REPORTED BENEFITS

Households or units (depending on the context) who are not classified as having imputed benefits. Unit composition may or may not have been imputed.

REPORTED INCOME AND BENEFITS

Households or units (depending on the context) in which no imputations were performed for either income amounts or food stamp benefits. Unit composition may or may not have been imputed.

ELIGIBILITY

An approximation of eligibility for the Food Stamp Program based solely on income, size and presence of elderly or disabled members. This was constructed for households using income, composition, and size of the Census household; and for units using income, composition, and size of the food stamp unit.

OUTLIERS

Households or units (depending on the context) with excessively high income.

7

B. OVERVIEW OF THE REPORT

The discussion is subdivided into three parts followed by a concluding chapter. Chapter II focuses on the economic characteristics of food stamp participants. The presentation begins with the distribution of households by income class as published by the Census Bureau. This table is then replicated several times with successively more restricting universe definitions to show the impact of the imputation process, i.e., the imputation of income, benefits or both, and the impact of the unit of analysis on this distribution. Following that, the economic focus switches to the eligibility concept which accounts for unit composition and size as well as income.

Chapter III shifts the focus from the distributional characteristics of the food stamp population to an analysis of the impact of the imputation of income or benefits on outliers. In this case outliers are households or units with benefits (reported or imputed) whose incomes (reported or imputed) are excessively high. The chapter further examines the relationship between average benefits and the determinants of benefit levels. This analysis, which is based on the food stamp unit, compares the relationship originally measured by the survey, i.e., that derived from units with reported income and benefits, to the relationship imposed on nonrespondents through the imputation of income or benefits.

Chapter IV describes the process employed in the construction of the food stamp unit for this analysis, focusing on the difficulty in interpretation of multiple units within the Census household. The difficulties described in that chapter result from the method of imputing unit composition, the lack of

8

a record of when unit composition was imputed, and the apparent duplicate reporting of benefits within the Census household.

C. SUMMARY OF FINDINGS

The number of food stamp households to which either income or food stamp benefits were imputed is a relatively small proportion of the sample with food stamps. As a result, the impact of the imputation of income and benefits on the aggregate distributional characteristics of the food stamp population is very small. When aggregate statistics are based on the food stamp unit rather than the Census household, there is a shift in the apparent economic status of food stamp participants. Units are somewhat poorer than households and there are considerably fewer units in the upper tail of the income distribution than there are households with food stamps. Examination of income relative to food stamp guidelines for eligibility shows that some food stamp households and some food stamp units do not appear to be eligible for the program. This finding is not unexpected. However, the proportion of the sample which does not appear eligible for the program drops by one half when the unit of analysis is changed from the Census household to the food stamp unit. In the aggregate the imputation of income or food stamp benefits does not have a dramatic effect on the proportion of households or units which do not appear to be eligible for the program.

The initial examination of aggregate statistics appears to show minimal effect of the imputation process on the distributional characteristics of food stamp recipients. However, as shown subsequently, this minimal effect is due more to the small proportion of the sample affected than to the imputation process

9

itself. Close examination of the food stamp households and units with imputed income or imputed benefits shows that the relationship between the benefits and the determinants of benefit levels is being distorted by the imputation of income and benefits. The imputation of income to food stamp units with reported benefits results in a disproportionately high number of units whose income exceeds the limits set for program eligibility and distorts the relationship between unit income, size, and benefit levels.

The imputation of food stamp benefits preserves the overall mean as expected. However, the distribution of imputed benefits by the determinants of benefit levels did not meet prior expectations based on the benefit formula. The variation in average benefits by the presence of elderly or disabled members for units with imputed benefits did not differ drastically from the expected variation based on reported benefits. This was attributed to the use of a close proxy for the presence of elderly in the imputation of benefits. A larger difference was observed for the variation in average benefits by the receipt of earnings, suggesting that the work experience variable used in the imputation process does not adequately account for the impact of earnings on benefit levels. The most extreme case was the variation in average benefits by poverty-level. As noted, imputed average benefits were approximately the same as reported average benefits. However, the group to whom benefits were imputed had higher incomes relative to poverty than the group with reported benefits, suggesting that their benefits should have actually been lower on average. This finding leads to the recommendation that a proxy for low income status be used in the imputation of food stamp benefits.

10

II.   ECONOMIC CHARACTERISTICS OF FOOD STAMP PARTICIPANTS

The U.S. Bureau of the Census (1985b) reported that 8000 households with average monthly income in excess of $6000 received food stamp benefits in the third quarter of 1984. This is a means-tested program and except for elderly and disabled households, units should not be granted benefits if their income exceeds 130 percent of the monthly federal poverty guidelines. Furthermore, the monthly poverty level for a household of size 4 in the third quarter is $850. Given that, how could there have been households with such high incomes participating in the program? These rather sensational statistics could reflect true circumstances in existence at that time. For example, poverty screens increase by unit size and hence extremely large units can have high incomes and still be eligible. This is particularly true for large units containing elderly or disabled members as these units are not subjected to the gross income test described earlier. Instead they are allowed to deduct certain expenses and if income less deductible expenses does not exceed 100 percent of poverty they are eligible (assuming their assets do not exceed $3000).

A second potential reason some food stamp households can have unusually high incomes is that the Census household is not precisely the same as the food stamp unit. Hence, the household's income is not a direct determinant of program benefits. Finally, it is possible that a household participating in the program early in the quarter could have had a windfall profit in the latter part of the quarter and then discontinued participation in the program.

Given the sensational nature of this statistic and the fact that it appears in an official government publication, we want to know--is it really true? We decided to address this issued by examining the same table for the month of August using Wave 4 of SIPP. August is the calendar month for which full panel estimates are obtainable with that wave. It also is a month for which administrative survey data have been collected on program participants and hence can be used for comparison purposes. The discussion of the published statistics from SIPP is subdivided into two parts. The first examines households and units by income class and the second examines households and units by an approximation of program eligibility status.

1. Households and Units by Income Class

Table 1 shows the distribution of households with food stamp benefits distributed by income class for both third quarter 1984 (from Table 9 of the Census publication) and for August (from tabulations of the Wave 4 microdata files). Of note is that the number of households in the highest income category in August is twice the number published by the Census Bureau. In addition to replicating the distribution of weighted cases, we have added to this table the number of unweighted cases on which the August statistics are based. We conclude from the calculation of average weights in the highest income class in August that the published reports on the number of food stamp households with income in excess of $6000 is based on less than 10 cases.[1] Furthermore,

---

[1] The Census Bureau used one third of the sample weights for each monthly observation to create average statistics for the calendar quarter. Assuming an average weight of 4000 per household (or 1333 for one third of the household weight) in the highest income category, the number of monthly observation on which the Census statistic is based is 6.

TABLE 1

DISTRIBUTION OF FOOD STAMP
HOUSEHOLDS BY MONTHLY CASH INCOME

Income Level	Third Quarter 1984[a] (1000's)	August 1984 Weighted[b] (1000's)	August 1984 Unweighted[b]
<300	1260	1266	256
300-599	2645	2667	567
600-899	1001	1015	224
900-1199	476	513	119
1200-1499	219	209	48
1500-1999	185	213	52
2000-2499	73	78	20
2500-2999	46	51	11
3000-3499	34	24	7
3500-3999	24	32	8
4000-4999	12	9	2
5000-5999	7	8	2
6000+	8	16	4
TOTAL	5990	6102	1320

[a]U.S. Bureau of the Census (1985b) Table 9.

[b]Special tables from Wave 4 of SIPP 1984 Panel.

13

it appears that 6 out of 13 cells defined by income levels in Table 9 of the Census report are based on fewer than 50 observations (these are the 6 highest income categories).

The Census Bureau acknowledges that these small estimates are not reliable noting that they were presented "primarily to permit such combinations of the categories as serve each user's needs." However, casual readers of the report in question are not likely to read that note of caution as it only appears in an appendix to the report.

Believing that such small sample sizes produce insignificant results, we will proceed with the analysis of table 9 collapsing the 7 high income groups into 1 cell representing monthly incomes of $2000 or more. Table 2 displays the distribution of food stamp households by income class altering the universe across the columns. The first universe is a replicate of the full set of households from Table 1. This is followed by the distribution for households with reported benefits, and finally the universe is restricted to those households with reported income and reported benefits.[1] Overall this comparison is encouraging because the benefit and income imputations performed for 12 percent of the unweighted sample cases did not drastically alter the distribution by income class.[2] It is useful to note at this time that in the original reported data, 27 cases out of 1158 (or 2.3%) fell into the high income group,

---

[1] We have not adjusted the weights for the restricted universe definitions.

[2] There are 1320 unweighted households with food stamp benefits, 162 or 12% of which had either some income or some food stamp benefits imputed.

14

TABLE 2

DISTRIBUTION OF FOOD STAMP
HOUSEHOLDS BY INCOME CLASS
August 1984

Income Class	All Households Counts	%	Households with Reported Benefits Counts	%	Households with Reported Income and Reported Benefits Count	%
<300	1266 (256)	21	1228 (248)	21	1202 (242)	22
300-599	2667 (567)	44	2579 (548)	45	2519 (534)	47
600-899	1015 (224)	17	933 (205)	16	844 (186)	16
900-1199	513 (119)	8	473 (109)	8	433 (100)	8
1200-1499	209 (48)	3	164 (37)	3	159 (35)	3
1500-1999	213 (52)	3	173 (43)	3	139 (34)	3
2000+	218 (54)	4	175 (42)	3	116 (27)	2
TOTAL	6102 (1320)	100%	5724 (1232)	100%	5412 (1158)	100%

Weighted counts in thousands. Unweighted counts in parenthesis. Totals may not add due to rounding.

SOURCE: Special tabulations from Wave 4 of SIPP 1984 Panel.

15

only one of which falls into the category of income in excess of $6000. On the other hand, after imputation for nonresponse, 54 out of 1320 or (4.1%) fell into the high income group, 4 of which now have income in excess of $6000 in the month of August. This suggests that the imputation process is producing too many outliers relative to reported data, an issue which is addressed in Chapter III.

The estimates discussed so far are based on households and income. It is useful to repeat the tables using the food stamp unit. This will demonstrate whether the high income households are concentrated among those with noncovered persons who received a large portion of the household's income.[1] The issues surrounding the existence of these individuals are not addressed here.[2] The concern is whether the publication of participants based on the Census household definition is misleading, especially for the easily identified program participants with high incomes.

Table 3 shows the distribution of food stamp units by income class for the three different universe definitions employed for Table 2 -- all units, units with reported benefits and units with reported benefits and reported income. A food stamp unit for this purpose is the group of household members covered

---

[1]As noted in the introduction, the food stamp unit definition is more restrictive then the Census household definition. Therefore, households receiving food stamps can legitimately contain persons who are not covered by the benefit.

[2]These issues are addressed in Landa (1987).

TABLE 3

DISTRIBUTION OF FOOD STAMP
UNITS BY INCOME CLASS
August 1984

Income Class	All Units		Units with Reported Benefits		Units with Reported Benefits and Reported Income	
	Counts	%	Counts	%	Counts	%
<300	1543 (329)	25	1496 (319)	26	1465 (312)	27
300-599	2934 (635)	48	2834 (613)	50	2764 (596)	50
600-899	898 (193)	15	814 (174)	14	748 (160)	14
900-1199	421 (93)	7	381 (83)	7	350 (76)	6
1200-1499	135 (29)	2	102 (21)	2	102 (21)	2
1500-1999	101 (23)	2	71 (16)	1	57 (12)	1
2000+	70 (18)	1	26 (6)	*	19 (4)	*
TOTAL	6102 (1320)	100%	5724 (1232)	100%	5504 (1181)	100%

Weighted counts in thousands.   Unweighted counts in parenthesis.   Totals may not add due to rounding.

*Less than .5%.

SOURCE: Special tabulations from Wave 4 of SIPP 1984 Panel.

17

by food stamps.[1] In 84 percent of the cases, the food stamp unit is identical to the household. For each of the remaining units (16 percent) at least one household member is not covered by the benefit.[2]

Due to the large proportion of cases where the household and the food stamp unit are the same, we do not expect to see large differences between tables 2 and 3 and in fact we do not. Overall, the imputation process affects 10 percent of the units as compared to 12 percent of the households, indicating that in 2 percent of the households, noncovered persons had some amount of income imputed and they were the only persons with income imputed.[3] As was true for households in Table 2, the imputation of benefits or income to 10 percent of the unweighted sample of food stamp units does not drastically alter the overall distribution by income class. Of particular interest, though, is the proportion of cases in the highest income category. In Table 2, 4 percent of all households (54 cases unweighted) had incomes in excess of $2000. This proportion declined to 2 percent (27 cases unweighted) when

---

[1]In households with multiple persons reporting food stamps, only one unit was constructed. This unit consists of all persons in the Census household who were covered under at least one of the recipients' benefits. See Chapter IV for further discussion of multiple-unit households.

[2]This compares favorably to the estimate of 18 percent obtained by Dalrymple and Carlson (1986) for September, 1983. Tabulations of Wave 3 show that 16 percent of Food Stamp households in April, 1984 contain a noncovered person.

[3]In other words, both Tables 2 and 3 have 1320 observations with Food Stamps. In Table 2 there are 162 households (1320 - 1158) with either benefits or income imputed. However, in Table 3 only 139 food stamp units (1320 - 1181) have these imputations. Thus 23 (or 2 percent) of the households have an imputation for some household member who is not part of the Food Stamp unit.

imputed households were eliminated. In Table 3, the proportion declined still further to 1 percent (18 cases) by simply narrowing the unit of observation to the food stamp unit. Finally, less than one half of one percent of all food stamp units with reported income and reported benefits had income in excess of $2000 (4 cases unweighted). Furthermore, none of those 4 unweighted cases fell into the category of $6000 or more in monthly income.[1]

The distribution by income class shifted somewhat between Tables 2 and 3 with more units than households in the lower classes (65 percent of all households had incomes below $600 whereas 73 percent of all units had this level of income). This shift in the distribution of cases to the lower income categories suggests that the food stamp unit's income is less than the household's income in at least half of the households containing a noncovered person.[2]

Table 4 demonstrates that the food stamp unit income is less than household income in most of the food stamp households with a noncovered person. Previously we reported that sixteen percent of food stamp households have at least one uncovered member, and Table 4 indicates that 15% of food stamp households contain a noncovered member receiving at least some of the household's income. The upper half of the table shows how households within each

_____

[1] Two had incomes in the 2000-2499 class and two had incomes in the 3500-3999 class.

[2] Food stamp unit income is either the same or less than household income depending on whether noncovered persons have any income. The shift in income distributions suggests that most noncovered persons do have income. The number of units with incomes less than $600 exceeds the number of households in that income category by 544 thousand, which represents 57 percent of the number of households with a noncovered person or 9% of all food stamp households.

TABLE 4

COMPARISON OF FOOD STAMP UNIT INCOME TO
HOUSEHOLD INCOME IN FOOD STAMP HOUSEHOLDS
August 1984

	Unit Income Less Than Household		Unit Income Equal to Household		Total	
	Count	%	Count	%	Count	%
HOUSEHOLD TOTAL INCOME						
<300[a]	151 (31)	12%	1115 (225)	88%	1266 (256)	100%
300-599	75 (18)	3%	2591 (549)	97%	2667 (567)	100%
600-899	189 (48)	19%	827 (176)	81%	1015 (224)	100%
900-1199	129 (36)	25%	384 (83)	75%	513 (119)	100%
1200+	341 (86)	53%	299 (68)	47%	640 (154)	100%
TOTAL	885 (219)	15%	5216 (1101)	85%	6102 (1320)	100%
UNIT INCOME AS PERCENT OF HOUSEHOLD INCOME						
0	206	3%	NA		NA	
1-25%	257	4%	NA		NA	
26-50%	219	4%	NA		NA	
51-99%	203	3%	NA		NA	
100%	NA		5216	85%	NA	
TOTAL	885	15%	5216	85%	6102	100%

Note: Weighted counts in thousands. Unweighted counts in parenthesis. Totals may not add to rounding.

Source: Special tabulation of SIPP Wave 4, 1984 Panel.

[a]There are 122 thousand food stamp households with no income. In all cases, the food stamp unit is the same as the household.

20

income class are distributed by whether or not a noncovered person receives at least some of the household's income. As expected, a higher percentage of the households with high incomes have uncovered members whose income is excluded from the food stamp unit than occurs in poorer households.

The lower half of the table describes how much of the household's income is being excluded from the food stamp unit. In roughly three-fourths of all households with a noncovered person (or 11 percent of all food stamp house-holds), the food stamp unit income is less than half of the household's income.

Although the comparison of tables 2 and 3 do not yield drastically different results in terms of the distribution of the food stamp population by income class, Table 4 suggests that presentation of the results on a household basis is misleading for the 16% of the food stamp households with an uncovered person.

2. Households and Units by Eligibility Status

The second case to be examined is one where the eligibility dimension is introduced. This will help to sort out apparent outliers from those house-holds with high incomes which might be entitled to benefits due to a large size or the presence of an elderly or disabled individual. It will also identify outliers of smaller size and less pronounced income that are not as noticeable in the aggregate distributions. For purposes of this discussion a full definition of eligibility has not been employed. Instead we are examin-ing income eligibility which can be measured with information obtainable from

the core questionnaire and therefore available for use in the imputation of food stamp benefits on the cross-section files.[1]

Before discussing the analysis of the distribution of cases by income eligibility status it is useful to note an important observation common among surveys of the population. There are cases where benefits are reported (or imputed) which are not consistent with observed (or imputed) income. Of particular concern is the existence of so-called seemingly ineligible participants. These are units which would be simulated to be ineligible for the Food Stamp Program if a model of the program eligibility determination process was executed on the microdata files. Czajka (1981) and Doyle, et al., (1986) report on their existence in SIPP, speculate about the reasons for their existence and note the problems they present in the analysis of participant characteristics and the population to which the Food Stamp Program is targeted.

The existence of seemingly ineligible participants is not just a phenomenon associated with general-purpose household surveys. Samples of administrative records of the Food Stamp Program also reveal a small number of such cases

---

[1]Units with no elderly or disabled members are income-eligible if gross income is less than 130 percent of monthly poverty (this is a requirement in the FSP regulations). Units with elderly or disabled members are classified as income-eligible for purposes of this study if gross income is less than 185 percent of poverty. (In the FSP regulations these units are not subjected to a gross income test. However, according to administrative data displayed in Appendix B, most elderly and disabled units have gross income under 185% of poverty). Gross income is similar to household total income. The difference is that earnings of students under 18 are excluded and net earnings rather than draw is used for self-employed persons.

(less than one half of one percent as displayed in Appendix B).[1] These cases may represent either errors in data collection (the contents of the administrative records are manually extracted from the case files) or errors in determining benefits for units applying for benefits.

In the context of a general-purpose household survey like SIPP, seemingly ineligible participants may represent true situations, e.g., they could be concealing some income from the food stamp case worker in the application for benefits but reporting true income to the SIPP interviewer, or they may be some of the small number of cases which were incorrectly processed in the Food Stamp Program. However, they could also be the result of some form of nonsampling error. Given that the true cause of the existence of seemingly ineligible participants is not actually known and that one potential factor is some form of nonsampling error, it does not seem reasonable to allow the imputation process to increase the relative occurrence of this phenomenon. We do expect to see an increase in the absolute number of seemingly ineligible participants given (as we will show) that some units are reported as such. However, in the absence of evidence to suggest that nonresponse is markedly greater among such units, we do not expect the imputation process to create a disproportionately higher number of seemingly ineligible units than seemingly eligible units.

---

[1]The survey of administrative case records was extracted from the Integrated Quality Control System (IQCS) and represents a sample of approximately 7000 cases participating in the Food Stamp Program in August 1984. The IQCS is described in Dalrymple and Carlson (1986).

Tables 5 and 6 display the results of Tables 2 and 3 with the added dimension of income eligibility. As noted, SIPP data do contain observations which appear to be ineligible for the program based on income. The proportion of households which appear ineligible for the program is affected somewhat by the imputation of income and benefits. Seven percent of the households with reported income and benefits appear ineligible for the program whereas nine percent of all food stamp households appear ineligible.

It is interesting to note that when food stamp units are examined there is a drop in the proportion of seemingly ineligible participants by about one half. The imputation of income and/or benefits affects this number to some extent but not drastically. It is also interesting that seemingly ineligible households and units are not concentrated in the high income categories. Almost half of the ineligible units have income under $1200 and roughly one fourth of the ineligible households have income under $1200. Similarly, there are some income-eligible households (3 percent) and units (2 percent) with monthly incomes in excess or $1200. In other words, the analysis of the distributions of households and units by income class presented earlier did not accurately reveal the extent to which outliers exist in the SIPP data. Specifically, there are small units with low incomes who do not appear eligible because their incomes exceed the allowable limits for their sizes. Similarly, there are large units with high incomes who appear to be potentially eligible because their incomes do fall within the allowable limits for their sizes.

TABLE 5

FOOD STAMP HOUSEHOLDS BY INCOME
ELIGIBILITY BY INCOME CLASS
August 1984

Eligibility[a] Income	All Households		Households With Reported Benefits		Households with Reported Benefits and Reported Income	
	Count	%	Count	%	Count	%
POTENTIALLY ELIGIBLE	5522 (1184)	91	5237 (1119)	91	5009 (1068)	93
<1200	5324 (1137)	87	5075 (1081)	89	4868 (1035)	90
1200+	198 (47)	3	161 (38)	3	141 (33)	3
NOT ELIGIBLE	579 (136)	9	487 (113)	9	403 (90)	7
<1200	137 (29)	2	137 (29)	2	129 (27)	2
1200-1999	257 (61)	4	201 (48)	4	174 (40)	3
2000+	185 (46)	3	149 (36)	3	99 (23)	2
TOTAL	6102 (1320)	100%	5724 (1232)	100%	5412 (1158)	100%

Weighted counts in thousands. Unweighted counts in parenthesis. Totals may not add due to rounding.

[a]POTENTIALLY ELIGIBLE if gross income <185% of poverty (for elderly and disabled) or if gross income <130% of poverty (for other households). Note that gross income differs slightly from household total income.

SOURCE: Special tabulations from Wave 4 of SIPP 1984 Panel.

TABLE 6

FOOD STAMP UNITS BY INCOME
ELIGIBILITY BY INCOME CLASS
August 1984

Eligibility[a] Income	All Units Count	%	Units With Reported Benefits Count	%	Units with Reported Income and Reported Benefits Count	%
POTENTIALLY ELIGIBLE	5796 (1251)	95	5497 (1183)	96	5306 (1140)	96
<1200	5677 (1225)	93	5406 (1164)	94	5216 (1123)	95
1200+	119 (26)	2	91 (19)	2	91 (19)	2
NOT ELIGIBLE	306 (69)	5	227 (49)	4	198 (41)	4
<1200	119 (25)	2	119 (25)	2	111 (23)	2
1200+	186 (44)	3	108 (24)	2	87 (18)	2
TOTAL	6102 (1320)	100%	5724 (1232)	100%	5504 (1181)	100%

Weighted counts in thousands. Unweighted counts in parenthesis. Totals may not add due to rounding.

[a]POTENTIALLY ELIGIBLE if gross income <185% of poverty (for elderly and disabled) or if gross income <130% of poverty (for other households). Note that gross income differs slightly from household total income.

SOURCE: Special tabulations of Wave 4 of SIPP 1984 Panel.

26

Earlier we noted that one household was reported to have income in excess of $6000 and to participate in the Food Stamp Program. Furthermore, the imputation of income and/or benefits added three additional households to this category. All of these households are seemingly ineligible (after accounting for income, size and presence of elderly or disabled members) hence eliminating the possibility that they represent unusually large households. There are, however, some households with unusually high monthly incomes (in the range $2000 to $4999) that appear to be eligible for the program based on income.

III. EFFECTS OF IMPUTATION ON OUTLIERS AND AVERAGE BENEFITS

In Chapter II we alluded to the possibility that the methods used to impute income and food stamp benefits produce too many outliers in the data as a result of the omission of income levels in the imputation of food stamp benefits and the omission of receipt of food stamps in the imputation of income. These omissions also led to some concern over the consistency between income and benefits. While it might be reasonable to see a shift in the distribution of the food stamp population by income class when responses are imputed (reflecting an assumption that nonrespondents are somehow different than respondents), we do not feel it is reasonable to see a change in average benefits for a specific poverty level, elderly/disabled, and earner classification when nonresponses are imputed. This is because benefits are an explicit function of these characteristics.

A. THE IMPUTATION PROCESS

Many users of SIPP have suffered from the lack of availability of adequate documentation on the procedures employed in the production of the SIPP micro-data files. As a result of this lack of documentation, this summary of the imputation process is less than complete but represents our best understanding of the imputation process.

When an individual has indicated that a particular means-tested transfer benefit has been received, the Census Bureau initiates an edit to fill in any missing responses on the unit composition variables and a hot deck procedure

28

to impute benefits when they are not reported.[1] The hot deck is a procedure whereby an individual with characteristics similar to the nonrespondent is located and then that individual's reported benefits are assigned to the nonrespondent. In the case of food stamps and other means-tested transfers the characteristics used to define "similar", are as follows:

    Sex
    Race
    Age
    Number of persons covered
    Work experience of the recipient.

Besides the omission of income already mentioned it is useful to explain that the age break is at 65. This is inconsistent with the key age break in the Food Stamp Program which is 60. As noted, attainment of age 60 of any member of the unit entitles the unit to be subjected to more liberal eligibility tests. Work experience of the recipient is related to but not quite the same as one of the key determinants of benefit levels which is the level of the unit's earnings.

When an individual indicates receipt of other income types but does not report the amount received, the Census Bureau employs a similar procedure to impute the missing income. The dimensions of the hot deck array vary somewhat from one income type to the other. For example, the following variables are used in the imputation of wages and salaries:

---

[1]This hot deck technique is similar to that employed for the Annual Demographic Supplement to the Current Population Survey which is described in Welniak and Coder (1980).

29

Occupation
        Sex
        Age
        Race
        Educational attainment.

The Census Bureau also employs imputation techniques when an individual within an otherwise successfully interviewed unit either cannot or refuses to respond to the entire questionnaire.[1] In this case a statistical match is performed and the responses of the linked respondent are imputed to the noninterview person. This is performed before the edit and imputation for item nonresponse.

The variables used to link nonrespondents to respondents in the match vary depending on whether or not the nonrespondent has prior wave data and the number of donor records in each cell defined by the interaction of the match variables. The minimum amount of information used in each match is age, sex, marital status and designated parent or guardian. Nonrespondents and respondents without prior wave data can also be matched on the following:

        Education
        Household relationship
        Veteran status

provided sufficient matchable donors exist. Nonrespondents and donors with prior wave data can be matched on those characteristics plus the following derived from prior waves:

        Income sources
        Asset sources.

———————————

[1]This statement does not apply to Wave 1 of the 1984 panel.

B. EFFECTS OF THE IMPUTATION PROCESS ON APPARENT OUTLIERS

Due to the small portion of the sample which is affected by the imputation process it is difficult to detect from the analysis in Chapter II whether the imputation process has an effect on the outliers. Therefore, this section examines the distribution of imputed cases along two dimensions to discern what the effects are. The first dimension is poverty status and the second dimension is income- eligibility status. These are not precisely the same concepts since the eligibility determination takes into account the presence of elderly or disabled members which has a significant effect on benefit levels.

Table 7 shows the distribution of all households with reported income by poverty level (column 1), as well as the distributions for food stamp households with reported benefits where those households with fully reported income (column 2) are separated from those households that have had some income imputed by the Census Bureau (column 3). This table demonstrates that imputed income for food stamp households (column 3) is more concentrated in the low percentage end of the poverty level scale then the distribution for all households (column 1). This is desirable for households participating in a means-tested transfer program. It appears that the variables used in the imputation process have controlled the process somewhat for a sensible assignment of income to food stamp households.

However, imputed income for food stamp households is not as concentrated in the low percentage end of the poverty level scale as food stamp households with reported income. The number of food stamp households reported to have

31

TABLE 7

DISTRIBUTION OF FOOD STAMP HOUSEHOLDS
BY POVERTY LEVEL BY INCOME IMPUTATION
August 1984

| Poverty Level | All Households with Reported Income | | Food Stamp Households With Reported Benefits | | | |
| | | | Reported Income | | Imputed Income | |
	Count	%	Count	%	Counts	%
≤50%	3798 (775)	6	1631 (337)	30	35 (8)	11
51-100%	6111 (1295)	9	2616 (562)	48	109 (24)	35
101-130%	4251 (898)	7	645 (138)	12	51 (12)	16
131-185%	7259 (1536)	11	291 (69)	5	36 (9)	11
≥186%	43605 (9152)	67	230 (52)	4	81 (21)	26
TOTAL	65023 (13656)	100	5412 (1158)	100	312 (74)	100

Weighted counts in thousands. Unweighted counts in parenthesis. Totals may not add due to rounding.

SOURCE: Special tabulations from Wave 4 of SIPP 1984 Panel.

income in excess of 130 percent of poverty is 9 percent of the number of households with reported income and benefits, almost half of which have income in excess of 185 percent of poverty. On the other hand, among households with reported benefits and imputed income the proportion in the two highest poverty classes increases to more than one third of the total. Most of these have income in excess of 185 percent of poverty. The cutoff at 185 percent of poverty is important because administrative sources show that very few elderly/disabled households exceed this limit (nonelderly/nondisabled units are not allowed to participate if gross income exceeds 130 percent of poverty).[1]

Another way of looking at the impact of the imputation process on outliers is to examine the distribution by eligibility status as displayed in Table 8. Only 7 percent of households with reported benefits and reported income appear to be ineligible for the program whereas 27 percent of households with reported benefits and imputed income appear ineligible. Hence, the imputation of income for food stamp households with reported benefits does disproportionately increase the size of the seemingly ineligible population.[2]

---

[1]Special Tabulations of the August 1984 extract from the IQCS (reproduced in Appendix B) show less than one half of one percent of food stamp units with elderly or disabled members have gross income in excess of 185% of poverty.

[2]We did not disaggregate the households by degree of imputation due to small sample sizes. As noted in Table 8, in two-thirds of the households with reported benefits and imputed income, the amount of imputed income exceeds 25% of total household income.

TABLE 8

DISTRIBUTION OF FOOD STAMP HOUSEHOLDS WITH
REPORTED BENEFITS BY ELIGIBILITY BY INCOME IMPUTATION
August 1984

Income Eligibility[a]	Reported Income		Imputed Income[b]	
	Count	%	Count	%
POTENTIALLY ELIGIBLE	5009 (1068)	93	227 (51)	73
NOT ELIGIBLE	403 (90)	7	85 (23)	27
TOTAL	5412 (1158)	100%	312 (74)	100%

Weighted counts in thousands. Unweighted counts in parenthesis. Totals may not add due to rounding.

[a]POTENTIALLY ELIGIBLE if gross income ≤185% of poverty (for elderly or disabled) and if gross income ≤130% of poverty (for other households). Note that gross income differs slightly from household total income.

[b]In two-thirds of the households with imputed income, the amount of income imputed exceeds 25% of total household income.

SOURCE: Special tabulations from Wave 4 of SIPP 1984 Panel.

We conjecture that this disproportionate effect on seemingly ineligible households can be avoided in the imputation of income to households with reported food stamp benefits without changing the basic method of imputation currently employed by the Census Bureau. A fairly simple approach would be to add the use of reported receipt of food stamp benefits (or some other indicator of low income status such as receipt of any means-tested transfer benefit) to the characteristics already in use for the imputation of income amounts when they are not reported.

C. AVERAGE BENEFITS

Benefit levels under the Food Stamp Program are calculated as an explicit function of selected unit characteristics, many of which are measured in the SIPP core questionnaire. Specifically, gross income is first computed for the unit (gross income is approximately equal to total income measured by SIPP). From this total, units are allowed to deduct 20 percent of earnings along with allowances for child care and shelter expenses up to a limit. Units containing an elderly or disabled member are allowed an additional deduction for medical expenses in excess of a threshold and are not subjected to a cap on the deductible shelter expenses. Gross income less the deductions noted is referred to as net income. Benefits are computed as the difference between the Thrifty Food Plan and 30 percent of net income. The Thrifty Food Plan takes into account the need for increased food allowances as the size of the unit increases.

As a result of this formula, benefit levels vary significantly according to the level of earnings, the presence of elderly or disabled members, income of

35

the unit and the unit's size.[1] This leads to the expectation that benefits derived from SIPP would vary in a predictable pattern along these dimensions. In particular, when income and composition are reported but food stamp benefits are not, the level of imputed benefits is expected to vary in a manner consistent with reported benefits along the dimensions noted above. Similarly, if benefits are reported but some or all of the income is missing, the level of imputed income should not distort the known relationships between income and benefits given a specific unit size and composition and presence of earnings.

1. Relationship Between Income and Benefits

Table 9 displays average benefits for food stamp units classified according to the type of imputation performed and contrasts those benefits with average benefits derived from a survey of administrative records for a comparable time period. The comparison to administrative data provides a point of reference in evaluating the reasonableness of the original reported data against which the imputed data are compared. The administrative data can also serve as a guide in our supposition about the unknown characteristics of the nonrespondents. Of course, any assumptions about their characteristics are necessarily weak since there are a number of factors which can contribute to a discrepancy between household survey data such as that derived from SIPP and administrative survey data. Nonresponse is only one of those factors.

_____

[1]Other factors such as expenses effect benefit levels as well, but these are not measured in the core part of the survey and hence are not candidates for use in the imputation of benefits.

TABLE 9

AVERAGE BENEFITS AND DISTRIBUTION OF
FOOD STAMP UNITS BY POVERTY LEVEL
August 1984

	Reported Income And Benefits	Reported Benefits And Imputed Income	Imputed Benefits		Administrative Data[1]
			All	Imputed Income	
AVERAGE BENEFITS	$114.61	$136.73	$114.47	$119.19	$114.69
DISTRIBUTION BY POVERTY LEVEL					
≤100%	84%	64%	50%	41%	93%
>100%	16%	36%	50%	59%	7%
UNWEIGHTED COUNTS	1181	51	88	61	6962

[1]Derived from special tabulations of the August 1984 extract of the IQCS.

Source: Special tabulations of Wave 4, 1984 panel of SIPP.

The food stamp units in Table 9 are subdivided into three main groups with the subset of the third group displayed separately. The three main groups are:

o   Reported benefits and reported income
o   Reported benefits and imputed income
o   Imputed benefits.

It should be noted that this classification is based on whether or not anyone in the unit had imputed income or benefits rather than any one in the Census household.[1] Because of the nature of this discussion, it is more appropriate to target the food stamp unit even though the Census Bureau can only realistically consider Census households in the imputation phase.

Overall the average benefits for units with reported benefits and income ($114.61) compare favorably with average benefits derived from administrative data ($114.69). Reported benefits for units who did not report all of their income, however, exceed that reported for the first group by 19 percent ($136.73 verses $114.61). Given the nature of the benefit formula, this difference in average reported benefits suggests that the income nonrespondents should be poorer on average than the respondents. Therefore, we expect that the group with reported benefits and imputed income should have a higher percentage of food stamp units below the poverty level than the group with reported income. However, only 64 percent of income nonrespondents were below

---

[1]As noted in the previous chapter, 2% of the food stamp households had all of their imputed incomes attributed to one or more noncovered persons. Units in these households are classified as having reported income.

poverty after the imputation of income whereas 84 percent of the income respondents were below poverty.

We have graphed the relationship between income as a percent of the poverty level and reported benefits for food stamp units to highlight how the imputation process has assigned too much income to units reporting a food stamp benefit but not all their income (Figure 1). Although the relationship between gross income as a percent of the poverty level and food stamp benefits is not strictly linear, we have represented the relationship as a linear approximation from the tabulations to simplify the discussion. The dashed line for the units with imputed income always lies above the solid line for units reporting their income. This means that the imputation process generally assigns incomes that are greater than the income reported by units with similar benefit levels. The graphed data points also indicate that the difference between reported and imputed income is larger for both very low and very high reported benefit levels than for benefit levels between $75 and $125 (approximately 21% of the weighted cases in this range.)

Returning to Table 9, average imputed benefits ($114.47) compare favorably to reported benefits as expected given that the imputation process preserves the mean. However, it is interesting to note that after imputation of income, units with imputed benefits have higher incomes on average (only 50 percent are below poverty) than either of the two groups with reported benefits. If the relationship between poverty level and benefits was preserved in the imputation process, we would not expect this to occur. Instead, with the same average benefits we would have expected to see a proportion of units in

FIGURE 1

Percent of Poverty Level
by Reported Benefit Level
for Food Stamp Units, August 1984

Poverty Level (in percent)

Reported Benefit Levels (in dollars)
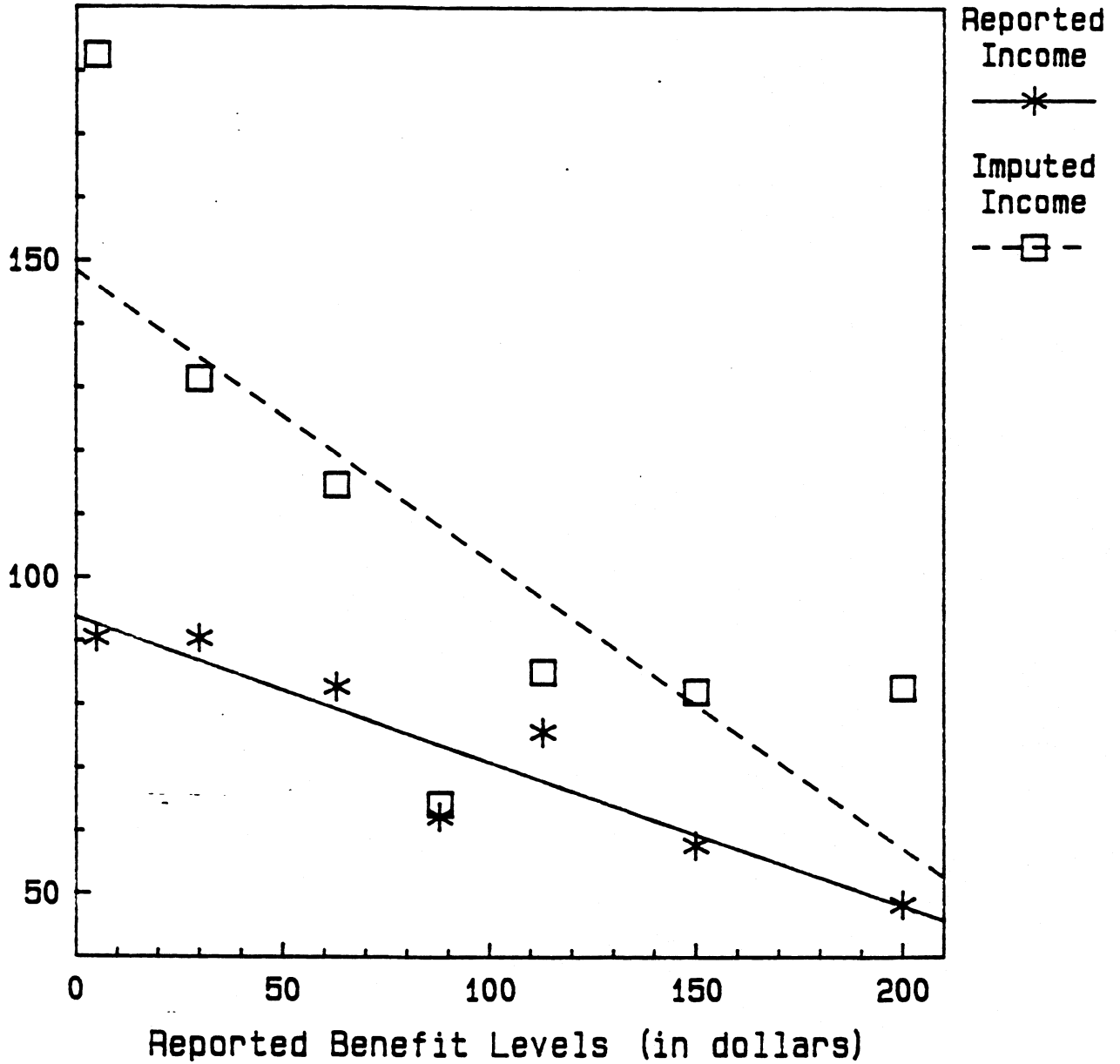
SOURCE: Special Tabulations from Wave 4 of SIPP
1984 Panel.

poverty comparable to the group with reported income and benefits. Of course, it is possible that the group with imputed benefits differs in some significant way from the group with reported income and benefits and this could explain a shift in the apparent income/benefit relationship. (For example, the group with imputed benefits may have higher expenses). However, we believe this is unlikely. Small sample sizes prohibited a cross classification along several dimensions needed to investigate this possibility.

The analysis of Table 9 suggests that while the imputation of benefits tends to preserve the overall mean, the relationship between income and benefits is being distorted. Similarly, the imputation of income to units with reported benefits does not maintain the expected relationship either.

This first point is amplified somewhat in Table 10 although few conclusive findings can be derived because the sample size is small. As expected, reported average benefits are higher for very poor households (under 50% poverty) than for the higher income groups. Average reported benefits are close to average benefits reported in the administrative data for this group as well. Reported average benefits decline for the higher income groups but this is not as rapid as the decline observed in the administrative data. In fact, reported average benefits for SIPP units above poverty are more than twice the average benefits of units above poverty based on administrative data. This phenomenon merits additional research.

Based on the deviation between reported benefits from SIPP and from administrative data and based on the observed fluctuation in average benefits by

TABLE 10

DISTRIBUTION OF AVERAGE FOOD STAMP BENEFITS
BY POVERTY LEVEL OF THE FOOD STAMP UNIT
August 1984

Poverty Level	Reported Benefits And Income	Imputed Benefits All	Imputed Income	Administrative Data[1]
≤50	162.08 (404)	101.70 (14)	95.97 (8)	169.74 (2749)
51-100	92.66 (598)	147.12 (28)	179.37 (15)	85.01 (3736)
101-130	79.73 (122)	89.30 (17)	89.02 (10)	32.40 (438)
131-185	86.43 (44)	83.75 (10)	80.67 (9)	11.10 (32)
186+	88.12 (13)	116.23 (19)	116.23 (19)	10.00 (7)
TOTAL	114.61 (1181)	114.47 (88)	119.19 (61)	114.69 (6962)

Source: Special tabulations of Wave 4 of 1984 panel of SIPP.

[1]Tables from the August 1984 IQCS extract.

poverty groups, several different suppositions can be made about the characteristics of nonrespondents and hence the level of missing benefits. First, given that reported benefits on average are less than administrative data for the very poor units one would expect that nonrespondents on average in the very poor group would have higher benefits then their reporting counterparts if the overall mean for the group is to match the administrative data. Correspondingly, given the reverse situation in the higher income groups one would expect that the rest of the nonresponding households would have lower benefits then their responding counterparts. On the other hand, if the original relationship noted in the SIPP data between benefits and poverty level were preserved, there would be little change in average benefits within poverty level between the reporters and the nonrespondents.

In actuality, however, none of these expectations has been met in the imputation of benefits. Imputed benefits for the very poor are less than reported benefits for that income class. Furthermore, among the other groups (with the exception of units in the 131-185% of poverty group) imputed benefits are larger on average than reported benefits.

Although small sample sizes prohibit any strong conclusions based on Table 10, the outcome depicted there does tend to reconfirm our earlier suspicions that the imputation process distorts the relationship between income and unit size and benefit levels. This was not observed in Chapter II due to the relatively small proportion of units affected and the focus on the aggregate distribution, not controlling for unit size. In fact, the relatively low level of nonresponse in SIPP is the reason this distortion does not seem to have a

43

large impact in SIPP-based statistics of aggregate caseload. However, we do think some attention should be given to this problem even though it only affects a relatively small number of observations. One suggestion for the prevention of at least some of the distortion is to perform imputations for missing food stamp benefits after the income has been imputed and to use some measure of the unit's poverty status (if unit composition was reported) or the household's poverty status (if unit composition was not reported) as a determinant of the level of imputed benefits. Similarly, if this type of distortion is observed to occur with other means tested benefits (and we assume it would) these benefits could be imputed in the same manner after other income amounts are imputed but before food stamp benefits are imputed.[1]

2. Presence of Elderly or Disabled

Table 11 displays the relationship between average benefits and the presence of elderly or disabled persons. As expected average benefits for units with reported income and benefits vary considerably by the presence of an elderly or disabled person. Average benefits for units with no elderly or disabled members exceeds those with an elderly or disabled member by a factor of 2.6. Although the reported average benefits from SIPP exceed that reported in the administrative survey for both groups this ratio of benefits between the two groups is the same.

---

[1]Food stamp imputations should come last because the eligibility and benefit determinations for food stamps include all forms of cash income in the measure of gross income.

TABLE 11

DISTRIBUTION OF AVERAGE FOOD STAMP BENEFITS
BY PRESENCE OF AN ELDERLY OR DISABLED
MEMBER IN THE FOOD STAMP UNIT
August 1984

Presence Of Elderly Or Disabled Member	Reported Benefits And Income	Imputed Benefits	Administrative Data[1]
NO	144.87 (780)	148.62 (44)	139.42 (4963)
YES	55.31 (401)	75.22 (44)	53.96 (1999)

Source: Special tabulations of SIPP wave 4 1984 panel.

[1]Tabulations of the August 1984 IQCS extract.

The imputation of benefits maintains the direction of the relationship (benefits for units with no elderly or disabled members exceed benefits for elderly/disabled units on average). However, the ratio of average benefits for the two groups decreases somewhat from 2.6 to 2.0. The use of an age break at 65 in the imputation of benefits appears to help to maintain this relationship. Lowering that age break to 60 (the food stamp designation of elderly) and incorporating the presence of elderly persons in the household other than the primary recipient might improve this result even more.

3. Presence of Earners

As noted previously, 20 percent of earnings is deducted in the computation of food stamp benefits. Therefore, given the same level of total income, a unit with earnings will have higher benefits than a unit without earnings. Administrative data confirm that units with earnings on average have higher benefits than units without (see Table 12). Benefits to earners exceed benefits to nonearners by about 19 percent. Units with reported benefits and income in SIPP also exhibit this characteristic, although the difference in benefit levels is somewhat larger (27 percent).

The imputation of benefits retains higher benefits for earners than for nonearners as expected given that at least some indication of earnings receipt is included in the imputation of food stamps. However, the ratio of the imputed benefits of earners to nonearners is 1.42, well above the 1.27 ratio for reported benefits. It is not clear why this ratio would increase so much. It does suggest that the current use of work experience of the

46

TABLE 12

DISTRIBUTION OF AVERAGE FOOD STAMP BENEFITS
BY PRESENCE OF EARNINGS IN THE FOOD STAMP UNIT

Presence Of Earnings	Reported Benefits And Income	Imputed Benefits	Administrative Data[1]
No	$107.21 (877)	$91.26 (50)	$110.72 (6433)
Yes	135.73 (304)	147.09 (38)	131.34 (1529)
TOTAL	114.61 (1181)	114.47 (88)	114.69 (6962)

Note: Unweighted Counts in parenthesis.

Source: Special tabulations of SIPP Wave 4 1984 Panel.

[1]Tabulations of August 1984 IQCS extract.

recipient in the imputation of benefits is not quite sufficient as a proxy for receipt of earnings by the food stamp unit.

Up to this point we have observed that estimates based on the food stamp unit do vary from estimates based on the Census household. Furthermore, we have based our comparisons to administrative data on the food stamp unit rather than the Census household. However, we have noted that some concern exists over the measurement of this unit, particularly with regard to households which contain more than one recipient of food stamp benefits. The next chapter elaborates our concerns and explains why our food stamp unit construct did not attempt to distinguish multiple units within Census households containing more than one recipient of benefits.

IV. CREATION OF FOOD STAMP UNITS

The food stamp unit definition as currently legislated allows for the formation of program units that are subsets of the Census Bureau's definition of a household. A Census household consists of persons who usually reside together at a specific address. The food stamp unit generally is the subset of those cohabitants who purchase and prepare meals together. There are exceptions to the general rule which permit multiple units to form even within groups of cohabitants who share meals. For example, an elderly or disabled person and his or her spouse are allowed to apply for benefits as a separate unit from the rest of the residents at their address.[1]

These regulations affecting the food stamp unit allow the formation of subunits and multiple units within participating Census households. In either case one or more individuals in the Census household may not be covered under food stamp benefits received by other residents at their address. Given the difference between program units and the Census household, one of the objectives of this paper was to discern whether the use of the Census household in the analysis of distributional characteristics of program participants produced different answers, particularly with regard to income and benefit levels, than would have resulted had the analysis been based on food stamp units. In order to accomplish this goal, food stamp units were constructed from the information available from the survey. This effort, described below,

---

[1]See Appendix C for a more indepth discussion of the definition of the food stamp unit and how that differs from the Census household definition.

49

was generally successful in identifying noncovered persons within Census households but not successful in identifying multiple units.

A. SIPP ENUMERATION OF FOOD STAMP UNITS

In the SIPP core module all persons over 18 and persons age 15-18 with dependents are asked if they were authorized to receive food stamp benefits some time during the four months prior to the interview date, i.e., did their name appear on the certification card. This method of questioning is intended to identify only one member of each food stamp unit in the Census household who would then be questioned further about their participation in the Food Stamp Program. Once the authorized recipients are identified, a series of questions designed to measure unit composition and monthly benefits are administered to those persons. Questions pertaining to benefit levels are administered four times, once for each of the four months prior to the month of interview. The unit composition questions are administered once per wave and inquire as to which persons "living here" were covered. There is some ambiguity in the responses to questions of unit composition when the Census household changed composition within the wave, especially when persons left the household during the reference period and before the interview month.

With this information, the Census Bureau creates microdata files reflecting food stamp benefits and unit composition in each of the reference months of the wave. Monthly benefits are recorded only on the authorized recipient's record whereas unit composition is recorded for all household members covered under someone's benefit. The unit composition is assigned in months in which benefits were received by setting a coverage flag to one for each person

50

covered. The determination of who was covered is accomplished in one of two ways. If the primary recipient indicated that all persons in the household were covered, then all persons residing at the recipient's address in the month or months in which benefits were received are listed as covered. If the primary recipient provided a list of covered persons (as they are requested to do when less than the full household was covered) then only those people listed are assigned coverage in the months in which benefits were received.

These monthly benefit and coverage fields are constructed on the public use microdata files after edits and imputations are performed. Imputation flags accompany the monthly benefit amounts indicating whether or not the benefit amounts in the microdata files had been imputed. The files do not contain flags to denote whether unit composition had been imputed. For this project these constructed variables were used to determine the food stamp unit within the Census household. Essentially the food stamp unit for purposes of this study consisted of all persons in the Census household whom the Census Bureau had flagged as being covered in the month of August 1984 in the SIPP Wave 4 1984 panel microdata file. Summary characteristics of the food stamp unit, such as income and size, represented aggregations over the covered individuals within participating households. Due to the existence of ambiguous cases described in the next section, only one food stamp unit was created in our analysis file when a Census household contained more than one authorized recipient. This unit consisted of all persons covered by at least one of the authorized recipients in the Census household. Food stamp benefits for the

51

unit represented the sum of all benefits reported by or imputed to the authorized recipients.

B. PROBLEMS IN THE FORMATION OF UNITS

The attempt to create "true" food stamp units for this paper met with a number of obstacles. These obstacles resulted from the method used by the Census Bureau to assign unit composition in cases of noninterview and in cases of nonresponse to the questions pertaining to food stamp benefits. In total, 214 thousand or 4 percent of the recipient households appeared to have been multiple food stamp units, i.e., there was more than one authorized recipient. This figure is consistent with the Dalrymple and Carlson (1986) estimate of 5 percent. However, upon close examination of the food stamp households with multiple recipients we found that the number of households with true multiple units may be as small as 1% of the total. There is considerable ambiguity in the data as illustrated below.

Food stamp households with more than one recipient were examined based on two criteria: (1) marital relationships between the food stamp recipients and (2) the potential that one or more recipient had unit composition imputed. Note that we did not know precisely who responded to the unit composition questions and who did not. The following situations were found to occur in 214 thousand households with more than one authorized recipient:

1. 28 percent or 59 thousand contained husband/wife families where both spouses reported being the authorized recipient.

2. 17 percent or 37 thousand were Census households where one of the authorized recipients was a noninterview and

52

was imputed food stamp receipt, benefits and unit composition where this unit composition duplicates what other respondents reported in the Census household.

3. 35 percent or 75 thousand of the Census households had food stamp benefits imputed (inclusive of the 17% noted above) and these units may or may not have had unit composition imputed.

Note that the duplicate reporting cases in (1) are not necessarily distinct from the cases with imputed benefits in (2) and (3).

Clearly, reporting of benefits by both spouses in a husband/wife family is suspect. Hence, if the 59 thousand duplicate reporting households are not counted as multiple units then the number of multiple unit households drops to 155 thousand or 2.5 percent of the total households with food stamps. It is possible that in some cases these represented true multiple units as in the case of a recent marriage of two single parents whose combined household had not yet been recertified as one unit. However, distinguishing a true multiple-unit husband/wife family from a situation of duplicate reporting of benefits is difficult with the cross-sectional data used for this task and hence was not attempted. Instead, we simply combined the units into one and summed the benefits reported by both spouses. The Census Bureau has been conducting research on this problem using longitudinal data and, we believe, they have been making an attempt to disentangle cases of duplicate reporting.

It is not clear whether true multiple units exist in food stamp households with multiple recipients with imputed benefits. For the third category listed above there is no method of ascertaining if unit composition was imputed in some fashion due to the lack of an appropriate imputation flag. We know that

53

in the 17 percent of the cases which contained a noninterview person (category 2) composition was imputed along with benefits, and that this resulted in duplicate coverage for at least some members of the household. Based on observation of the raw cases it appeared that composition was imputed along with benefits in at least some of the remaining cases in category 3. If we assume that all cases in category 3 had composition imputed along with benefit levels and that the resulting multiple recipients did not represent true multiple units, then the number of true multiple unit cases reported drops from 4 percent to 1 percent of the total caseload. If, on the other hand, we assume that none of these cases had unit composition imputed along with benefits and hence they represented true multiple units, then the number of multiple-unit households is reduced by 37 thousand.

Since imputation often creates duplicate coverage of Census household members, we do not believe that all of the imputed cases represented true multiple units. Hence, it appears that a sizeable portion of what seemed to be multiple-unit households were artificially created either through duplicate reporting of benefits or through the imputation process. We conjecture that the number of true multiple units is in the range of 86 thousand (or 1 percent of the total number of households with food stamp benefits) to 114 thousand (or 2 percent of the total)[1]. Given the relatively small size of the group of

---

[1]These estimates take into account the overlap between the three categories. Both figures assume that duplicative reporting (category 1) and imputation for noninterview (category 2) do not reflect the multiple units. The lower figure assumes that all the cases belonging exclusively to category 3 are not true multiple units. The higher figure assumes that they are true multiple units.

apparent multiple units and this ambiguity in distinguishing among separate units within a group of covered individuals, the food stamp unit file which formed the basis of many of the tables in the body of this report contained only one record per recipient household. Furthermore, that record pertained to all covered persons regardless of whether there appeared to be multiple recipients of food stamp benefits.

V. CONCLUSION

We have noted that the series of quarterly statistics on income and program participation originally published by the Census Bureau contained some unrealistic and unreliable estimates of the distribution of food stamp households by income class. Although the publication included some insignificant figures, we are not critical of the since-discontinued quarterly data series as a whole. In fact, that data series with a revised format would continue to be a useful reference. The format revisions would redefine the classification of food stamp cases to avoid publication of unreliable statistics and change the reference period to a calendar month. The latter change would have the benefit of providing some desperately needed controls for users of the public use microdata products.

This report addressed the impact of the imputation process on the relationship between the direct determinants of food stamp benefit levels--income, unit size, unit composition and earnings--and the benefit levels themselves. Heeringa and Lepkowski (1986) note that imputation for a missing item may distort the relationship between it and other items in the data unless specific controls are imposed. We have demonstrated this to be the case in examining the relationship between benefits and the determinants of benefit levels in SIPP. The lack of attention to reported food stamp benefits in the imputation of income appears to distort the observed relationship between poverty level and benefits for the six percent of the unweighted sample of food stamp households who had reported benefits and missing income. The first

56

finding was that the imputation of income to households with reported food stamp benefits produced a disproportionately high number of outliers which were defined in two ways. First, outliers were households whose income was in excess of 185 percent of poverty and second they were defined as potentially ineligible based on income. In both cases the proportion of outliers among food stamp households with imputed income and reported benefits was more than 3 times higher than the proportion of outliers among food stamp households with reported income and benefits (Food Stamp Households with imputed income and reported benefits comprised 6 percent of the unweighted sample of food stamp households whereas food stamp households with reported income and benefits comprise 88 percent).

This finding was further supported by an examination of the relationship between average benefits and total income for food stamp units. Specifically we found that reported benefits were higher for the group of households with missing income data than for households with fully reported income. This suggests that the income nonresponding households should be poorer on average given the inverse relationship between income and food stamp benefits. However, after imputation for the missing income amounts, this group actually had more income. Similarly, we found that although the imputation of benefits preserved the mean, the average imputed benefits were not realistic when examined separately for units grouped by poverty level. We believe that further dimensioning the Census imputation process for the relationship between income and means-tested benefits would very likely reduce the number of outliers currently produced by the imputation process.

We did find that the specific relationship between imputed benefits and the presence of an elderly or disabled member was more reasonable than the general relationship between imputed income and reported or imputed benefits. This was attributed to the use of some measure of the presence of an elderly member in the imputation of benefits. The specific relationship between the presence of earners and average benefits was not quite as reasonable leaving us with the belief that the imputation of benefits could be improved if a better proxy for earnings receipt was used in the imputation of benefits.

The fourth chapter of this paper described the construction of food stamp units within Census households. In attempting to create these units a number of problems arose with the determination of unit composition. Specifically, in an unexpectedly large number of cases, husbands and wives both reported being the authorized recipient of food stamp benefits. We also observed that the imputations for noninterview produced instances of duplicate coverage of some household members. Finally, it appeared that assignment of unit composition in situations where unit composition was missing resulted in duplicate coverage. As noted, however, this cannot be confirmed because the public use files do not contain a record of the imputation of unit composition.

REFERENCES

Czajka, J.L. Determinants of Participation in the Food Stamp Program: Spring 1979. Report to U.S. Department of Agriculture. Washington, D.C.: Mathematica Policy Research, Inc., November 1981.

Dalrymple, R. and S. Carlson. "Food Stamp Participation: A Comparison of SIPP with Administrative Records," Proceedings of the Bureau of the Census Second Annual Research Conference. Washington, D.C.: U.S. Bureau of the Census, 1986.

Doyle, P., J. Czajka, P. Boldin, H. Beebout and S. Hirabayashi. "Conceptual Studies in Microsimulation Using SIPP." Draft Report to the Food and Nutrition Service, USDA. Washington, D.C.: Mathematica Policy Research, Inc. November 1986.

Heeringa, S.G. and J.M. Lepkowski, "Longitudinal Imputation for the SIPP". Presentation at the American Statistical Association Joint Statistical Meetings, Chicago, Illinois. August 1986.

Kasprzyk, D. "A Review of the Research and Design Issues in the Survey of Income and Program Participation," Proceedings of the Survey Research Methods Section, American Statistical Association. Washington, D.C.: American Statistical Association, 1986.

Landa, C. "Exploring Hypotheses to Explain the Formation of Food Stamp Units in Households." Draft Report to the Food and Nutrition Service, USDA. Washington, D.C.: Mathematica Policy Research, February, 1987.

Little, R.J.A., "Missing Data in Census Bureau Surveys." Proceedings of the Bureau of the Census Second Annual Research Conference. Washington, D.C.: U.S. Bureau of the Census, 1986.

Nelson, P., D.B. McMillen and D. Kasprzyk. An Overview of the Survey of Income and Program Participation: Update 1. SIPP Working Paper Series 8401. Washington, D.C.: U.S. Bureau of the Census, 1985.

U.S. Bureau of the Census. Economic Characteristics of Households in the United States: Second Quarter, 1984. Current Population Reports, Series P-70, No. 4. Washington, D.C.: U.S. Government Printing Office, 1985.

U.S. Bureau of the Census. Economic Characteristics of Households in the United States: Third Quarter, 1984. Current Population Reports, Series P-70, No. 5. Washington, D.C.: U.S. Government Printing Office, 1985.

Welniak, E.G., and J.F. Coder. "A Measure of the Bias in the March CPS Earnings Imputation Scheme". Proceedings of the Survey Research Methods Section, America Statistical Association. Washington, D.C.: American Statistical Association, 1980.

APPENDIX A:

SELECTED TABLES FOR APRIL 1984 DERIVED FROM SIPP WAVE 3

TABLE A-1

DISTRIBUTION OF HOUSEHOLDS BY MONTHLY CASH INCOME

Income Level	Second Quarter 1984[a] (1000's)	April 1984 Weighted[b] (1000's)	April 1984 Unweighted[b]
<300	1311	1335	294
300-599	2779	2882	634
600-899	1059	1029	228
900-1199	460	472	110
1200-1499	218	198	48
1500-1999	220	220	51
2000-2499	94	80	22
2500-2999	55	64	17
3000-3499	39	48	12
3500-3999	26	15	4
4000-4999	24	18	5
5000-5999	3	-	-
6000+	4	4	1
TOTAL	6292	6364	1426

[a]U.S. Bureau of the Census (1985a) Table 9.

[b]Special tabulations from Wave 4 of SIPP 1984 Panel.

TABLE A-2

DISTRIBUTION OF HOUSEHOLDS WITH FOOD STAMPS BY INCOME CLASS
April 1984

Income Class	All Households		Households with Reported Benefits		Households with Reported Income and Reported Benefits	
	Count	%	Count	%	Count	%
<300	1,335 (294)	21	1,286 (283)	22	1,258 (277)	23
300-599	2,882 (634)	45	2,759 (606)	46	2,618 (576)	48
600-899	1,029 (228)	16	963 (215)	16	855 (189)	16
900-1199	472 (110)	7	413 (96)	7	337 (79)	6
1200-1499	198 (48)	3	169 (40)	3	105 (24)	2
1500-1999	220 (51)	3	196 (44)	3	123 (27)	2
2000+	229 (61)	4	194 (49)	3	135 (32)	2
TOTAL	6,364 (1,426)	100%	5,980 (1,333)	100%	5,430 (1,204)	100%

Weighted counts in thousands. Unweighted counts in parenthesis. Totals may not add due to rounding.

SOURCE: Special tabulations from Wave 3 of SIPP 1984 Panel.

DISTRIBUTION OF FOOD STAMP UNITS ON THE AUGUST 1984 EXTRACT
FROM THE INTEGRATED QUALITY CONTROL SYSTEM

Gross Income as Percent of Poverty	Total		No Elderly or Disabled		Elderly or Disabled Member	
	%	Count	%	Count	%	Count
≤ 50%	40	2873322	51	2640487	11	232835
51-100%	54	3900579	44	2297491	76	1603088
101-130%	6	440526	4	221980	10	218546
131-185%	1	51396	*	8011	2	43385
186%	*	7859	*	1108	*	6751
ALL	100	7273682	100	5169077	100	2104605

Totals may not add due to rounding.   Unweighted sample cases = 6917.

*Less than .5%

APPENDIX C:

SUMMARY OF FOOD STAMP PROGRAM UNIT RULES

APPENDIX C: SUMMARY OF FOOD STAMP PROGRAM UNIT RULES[1]

Food Stamp Program rules specify that the food stamp unit is to consist of all persons living together who customarily buy their food and prepare meals as a unit. Thus, it is possible for food stamp units consisting of persons who customarily buy food and eat meals apart from others in the Census household to legitimately form under Food Stamp Program rules.

There are certain exceptions to this general rule. Some related persons who live together must be members of a common food stamp unit, even if they do not customarily buy food and prepare meals together. Parents who reside with their children must be members of the food stamp unit to which their children belong, if one of the parents is not elderly or disabled. Children under 18 who live with their parents must be members of the food stamp unit to which their parents belong. Spouses who live together must be jointly covered; and siblings who live together, if one is not elderly or disabled, must be members of the same food stamp unit. In addition, persons who are elderly and disabled, and their spouses, can form a separate food stamp unit even if they buy food and prepare meals in common with other household members, as long as

---

[1]This appendix is an attempt to summarize six pages of federal regulations which are extremely complex. As such, it may differ in some respects from the regulations themselves. In these cases, the regulations and not the statements in this text are the official rules. The regulations governing the formation of food stamp units are 7 Code Federal Register 273.1.

the gross income of all other persons with whom they reside does not exceed 165 percent of the poverty line.[1]

Generally, residents of institutions are ineligible for food stamps but the law states that certain group living situations are not considered institutions. Residents of federally subsidized housing for the elderly, certain disabled or blind recipients of Social Security assistance in small group living situations, temporary residents of nonprofit shelters for battered women and children, and addicts or alcoholics in certain residential treatment programs are not considered residents of an institution. Such individuals are considered households for food stamp purposes (i.e. food stamp units).

Elderly or Disabled Persons in the Household

The presence of elderly or disabled persons in a food stamp household does not guarantee the existence of a food stamp unit consisting of a subset of the residence in the dwelling but it does increase the likelihood that such food stamp units will be formed.

---

[1] In addition, there are certain Census household members who cannot be covered by food stamp benefits, including: ineligible aliens, persons who fail to provide social security numbers, person who have intentionally violated Food Stamp Program rules, persons violating employment and training program or workfare rules, certain college students who are between 18 and 60 years of age, and SSI recipients in cash-out states. Further, roomers, and live-in attendants cannot belong to food stamp units with others in the household, but they can form a food stamp unit of their own. Finally, a boarder can be considered a member of a household to whom he or she is paying board at the household's request, but cannot form a separate food stamp unit. Also, boarders in commercial boarding houses are not eligible for food stamps.

As noted above, persons who are elderly and disabled and their spouses can form a separate food stamp unit even if they eat together with others in the household because they are unable to prepare meals on their own as long as the gross income of others in the household does not exceed 165 percent of the poverty line. In addition, households which contain elderly or disabled persons who live with their adult children or a sibling can form food stamp units if the elderly or disabled persons buy their food or prepare their meals apart from other members of the household. In the former case, one of the units must consist of the elderly or disabled person and his or her spouse. In the latter case, the food stamp unit may not necessarily contain the elderly or disabled member, it may be that only residual household members are covered under food stamps.

It is important to note that it is possible that households with elderly or disabled persons may not form food a stamp unit consisting of a subset of the residents. Unless there is a person in the household who is elderly and disabled, a household which has elderly and/or disabled persons cannot have uncovered members or multiple units unless there are persons who buy food and prepare meals separately. Even if the elderly or disabled persons in the household buy food or prepare meals apart from others in the household, they must be jointly covered with their spouse or children under the age of 18 if the spouse or children are present in the household. If the household consists entirely of an elderly or disabled person plus that person's spouse or under-18 children, at most one food stamp unit can legitimately exist and that unit must be the entire household. Even elderly and disabled persons

71

must be covered with their spouses; so that if the household consists solely of an elderly and disabled person and that person's spouse, at most one food stamp unit can legitimately exist and that unit must be the entire household.