Survey of Income and Program Participation

INVESTIGATION OF POSSIBLE CAUSES OF
TRANSITION PATTERNS FROM SIPP

by

Lynn Weidman
Bureau of the Census

No. 8713 35

November 1987

TABLE OF CONTENTS

INVESTIGATION OF POSSIBLE CAUSES OF TRANSITION PATTERNS FROM SIPP

I. Executive Summary

The purpose of this work is to look for possible causes for the patterns of between/within wave transitions reported by Burkhead and Coder (1985). They found that much larger proportions of transitions between receipt and nonreceipt of several income sources and food stamps were reported as occurring between waves than within waves. In this investigation we found no major influences from demographic characteristics, interview status, imputation procedures or longitudinal editing on this reported pattern. However, some smaller scale effects were noted.

1. For food stamps and social security, larger proportions of receipt of sources were reported by self-respondents than by proxies. There is usually a higher proportion of transitions between waves when at least one of two consecutive months has a proxy response than when both of the months are self-reported. (See Appendix 1)

2. There is a larger proportion of between wave transitions when at least one of two consecutive months is imputed than when both of the months are reported. For unemployment compensation the same is true for within wave transitions, where almost all imputes are needed when both of the months require imputation. (See Appendix 2)

3. For state and supplemental unemployment compensation a small but noticeable increase in the ratio of between to within wave transitions results from reimputation by the longitudinal edit. (See Figure 3.8)

There are certainly other relationships of small magnitude that will be found by further examination of the data, but nothing on the order of magnitude of the between/within wave differences. Further work should concentrate on improving the quality of responses.

II. Description of Study

The paper by Burkhead and Coder (1985) prepared for the ASA annual meeting presented an examination of reporting patterns of receipt and nonreceipt for income sources from the first twelve months of the 1984 SIPP Panel. For two consecutive months there are four possible receipt states: RR, RN, NR, and NN, where R=receipt and N=nonreceipt. They found that it was not unusual for two to three times as many changes between receipt and nonreceipt, i.e, transitions, to be reported between the last month of one wave and the first month of the next wave as were reported between any two consecutive months of either wave. This pattern would make it difficult to identify characteristics such as short term changes in the proportion of people receiving a given income source and the average length of spells of receipt and nonreceipt.

In order to make these data more usable we would like to be able to determine causes for the observed patterns and either adjust the data or change the questionnaire and procedures accordingly. It has been hypothesized that 1) respondents who tend to report transitions between waves can be identified by demographic characteristics or 2) changes between proxy and self-respondents from wave to wave introduce transitions. Distributions of transition counts were calculated in order to examine the possibility of the existence of four effects.

1. Effect of demographics on transitions.
2. Effect of interview status on transitions.
3. Effect of imputation on transitions.
4. Effect of longitudinal edit on transitions.

Food stamps were examined along with seven income sources: social security, unemployment conpensation, private pensions, VA compensations and pensions, supplemental security income, child support and AFDC. These were selected because several analysts agreed that they were the most important sources to explore initially.

A longitudinal file which includes demographic variables along with recipiency information and amounts for these income sources was created by the Population Division. This file of monthly data is based on the first four waves of data available from each household. Each of these waves is searched for all persons who reported receipt of any of the income types of interest during any month of the wave. For each such person all the information available for the 16 month period is collected and placed on a record.

This record was then used for computations if the person was interviewed for each of the four waves. All computations are performed on the contents of this file. For the study of (1) and (2) we looked at the distributions of receipt states using their total frequencies between all three wave pairs and within waves 2, 3 and 4. (A more complete description of the data used is given in Weidman (1986).) Therefore this work refers only to a small proportion of SIPP respondents. This means that proportions based on total responses are much larger than they would be if we used a denominator based on all the respondents, and this should be kept in mind when examining the tables in the appendices. However, all transitions are included on this file.

It is important to note that no statistical tests were performed in this study. In the executive summary and later in this report comparisons of proportions are made by examining distributions. Statements such as "a larger proportion" refer to a noticeable difference that would probably be identified as "significant" by an appropriate statistical test.

III. Demographics and Transitions

The intent of this work was to determine if the respondents that reported transitions between waves had different demographic characteristics than those that reported transitions within waves. A detailed description of this work is given in Weidman (1986).

The demographic variables and the categories used are defined as follows:

age: 15-30, 31-45, 46-60, 61+
sex: male, female
race: white, nonwhite
education: elementary, high school, above high school
marital status: married, (separated, divorced, widowed), never married
household size: 1,2,3,4-5,6+
tenure: home owned, not owned
relationship to reference person: reference person, spouse, child, other
SMSA size: not in SMSA, less than 1 million, 1 million +

For individual income types there are some differences between demographic groups, but none on the order of the size to indicate primary causation of the between/within wave differences. As an example, see Tables 3.A and 3.B in Weidman (1986).

IV. Interview Status and Transitions

For each person included in a SIPP household there are two possible interview statuses each month: S=self and P=proxy. We will refer to the combination of interview statuses for two consecutive months as an interview state. Within waves there are two possible states--SS, and PP where the proxy is the same person for both months. Between waves there are four possible interview states -- SS, SP, PS and PP. In this case the two proxies may be the same person or two different persons. Appendix 1 includes tables of distributions of transition state by interviewer state. The between wave frequencies are totals of receipt states reported in months 5, 9 and 13. The within wave frequencies are totals over months 6-8, 10-12 and 14-16. The total of frequencies within waves is exactly three times the total for between waves.

It was suspected that the reported percentage of transitions would be much higher for interview states that involved proxies than for SS. Food stamps, social security and unemployment compensation (Tables 1.1, 1.4 and 1.8) were the sources with relatively large numbers of transitions reported (i.e., with enough transitions to compare distributions for many cells.) The first two of these sources showed about the same patterns: (i) the proportion of transitions between waves was usually 1.5 to 2.5 times larger for SP, PS, and PP than for SS; (ii) the proportion within waves was about the same for SS and PP; (iii) the ratio RR/NN within waves is smaller for PP than for SS and between waves is usually smaller for PP, PS, and PP than for SS. Pattern (iii) is much more pronounced for food stamps than for social security and shows that larger proportions of receipt of sources were reported by self respondents than by proxies. Because the number of SS cases was much larger than the sum of SP, PS, and PP cases, these patterns did not have a noticeable effect on the within/between wave differences.

V. Imputation and Transitions

Is it possible that the method of imputation caused transitions to occur that would not have been reported? (Note that nobody can determine from the analysis file whether the receipt state and the amount or just the amount has been imputed. Thus we say imputation "is involved" with a transition to indicate that a change in receipt state was either (a) imputed or (b) reported and the amount was imputed.) For each of the sources there are a number of between wave transitions with imputation involved, ranging from .09 to .26 of the number reported. Excluding unemployment compensation, there are only 8 cases in which imputation was involved within a wave. Thus the proportions of transitions involving imputation are much larger between waves. Tables comparing transitions with and without imputations involved are given in Appendix 2. Again the between wave frequencies are totals over all transitions reported in months 5, 9 and 13. Within have frequencies are totals over months 2-4, 6-8, 10-12 and 14-16. The total of frequencies within waves is approximately four times the total for between waves.

For all sources the proportion of between wave receipt states that are transitions are larger when imputations are involved than when they are not. This result indicates that further investigation into the relative frequency of recipiency and amount imputation for the first and last months of a wave would be useful. If amounts are usually being imputed, then the transitions actually occur but the amounts are unknown. If receipt state is being imputed quite often, then this says something about the imputation mechanism or the characteristics of the people who require imputation.

Unemployment compensation has a higher proportion of transitions than any other source and is the only one with a large enough number of imputed within wave transitions to investigate. Almost all within wave imputations occur when both months involve

imputation. Does this mean that people requiring within wave imputation of unemployment compensation tend to be more like those people reporting transitions than those who do not? Or is there some other mechanism causing this pattern? These questions should be answered.

VI. Longitudinal Edit and Transitions

Population Division provided the longitudinal file before and after it had been through the "longitudinal" edit. This allowed an examination of changes in receipt states and transitions created by this edit. Appendix 3 contains this information for the eight income sources. Each income source has four tables: (1) receipt state totals reported between waves from the data before longitudinal edit; (2) receipt state totals for each within wave month before longitudinal edit (3) change in between wave receipt state totals after longitudinal edit (- means a reduction in the total after the edit); and (4) change in within wave receipt state totals after longitudinal edit.

Before the longitudinal edit, the maximum reported proportion of transitions within waves for any source but unemployment compensation is .07. For unemployment compensation their proportion is .21. In no case is the number of records in state NR or RN increased by the edit. There is a pattern of change for RR and NN only for unemployment compensation, where RR increases for each month and NN decreases. This same pattern holds for NN and RR between waves. However, only for NR in month 5 is there a decrease. Overall, there are 1.072 times as many transitions after the longitudinal edit as before between waves and only .912 times as many within waves. Thus the longitudinal edit somewhat increases the between/within wave transition discrepancy for unemployment compensation.

VII. Discussion

None of the computations indicated effects large enough to contribute markedly to the between/within wave pattern of reported transitions. In the demographic analysis the distributions of receipt state were examined for nine characteristics and all 72 pairwise combinations of them. The absence of any notable relationships seems quite revealing.

Interview status and imputation are other possible sources of differences and the results for them were similar. A couple of points needing further investigation were indicated. These results strongly suggest that the reason for the between/within wave pattern is that it is often simply easier to give the same responses for all four months of an interview than to recall in detail monthly transitions and amounts.

It seems that further work on the causes of transitions for these income sources and food stamps would yield little additional information. The pattern of response has been determined and we will have to live with it unless changes are made on the questionnaire or the way in which questions are asked. This means that we will not be able to pinpoint the exact time at which changes in the level of receipt occur, but must settle for being close. The estimation of lengths of recipiency spells will also be affected.

Further work in this area should concentrate on the survey instrument and its administration in order to improve the accuracy of transition reporting. At the same time we should make an effort to determine how our estimates are affected by the current reporting pattern and what adjustments can be made to improve these estimates.

References

Burkhead, Dan and John Coder (1985). "Gross Changes in Income Recipiency From the Survey of Income and Program Participation." American Statistical Association, Proceedings of the Social Statistics Section.

Weidman, Lynn (1986). "Investigation of Gross Changes in Income Recipiency From the Survey of Income and Program Participation." American Statistical Association, Proceedings of the Survey Research Methods Section.

APPENDIX 1

Transition State Distribution by Interview State

Table 1.1

Social Security

Receipt State	Interviewer State: Between Waves SS	SP	PS	PP	Within Waves SS	PP
RR	.931 (13805)	.906 (1490)	.854 (1139)	.862 (2369)	.942 (45713)	.907 (11958)
RN	.007 (104)	.015 (24)	.016 (21)	.017 (46)	.002 (78)	.003 (41)
NR	.015 (219)	.027 (44)	.030 (40)	.100 (89)	.005 (251)	.005 (60)
NN	.048 (707)	.053 (87)	.032 (134)	.089 (244)	.051 (2465)	.085 (1120)

first number in cell is proportion of column total
second number in cell is frequency count

Table 1.2

Veterans Compensation

Receipt State	Interviewer State: Between Waves SS	SP	PS	PP	Within Waves SS	PP
RR	.904 (1263)	.921 (174)	.88 (147)	.92 (381)	.919 (4312)	.941 (1703)
RN	.016 (22)	.005 (1)	.03 (5)	.01 (4)	.003 (13)	.001 (2)
NR	.021 (29)	.021 (4)	.024 (4)	.022 (9)	.003 (13)	.001 (1)
NN	.059 (83)	.053 (10)	.066 (11)	.048 (20)	.075 (354)	.057 (103)

Table 1.3

AFDC

Receipt State	Interviewer State: Between Waves SS	SP	PS	PP	Within Waves SS	PP
RR	.692 (1273)	.607 (116)	.556 (99)	.488 (101)	.731 (4424)	.604 (721)
RN	.061 (113)	.105 (20)	.034 (6)	.077 (16)	.016 (96)	.008 (10)
NR	.058 (107)	.042 (8)	.112 (20)	.097 (20)	.016 (98)	.008 (10)
NN	.188 (346)	.246 (47)	.298 (53)	.338 (70)	.237 (1433)	.379 (453)

Table 1.4

Food Stamps

Receipt State	Interviewer State: Between Waves SS	SP	PS	PP	Within Waves SS	PP
RR	.634 (2782)	.498 (231)	.409 (179)	.425 (263)	.658 (9532)	.509 (1653)
RN	.073 (319)	.108 (50)	.123 (54)	.121 (75)	.017 (245)	.02 (66)
NR	.053 (234)	.086 (40)	.087 (38)	.06 (37)	.022 (313)	.022 (71)
NN	.24 (1054)	.308 (143)	.381 (167)	.394 (244)	.303 (4391)	.449 (1459)

Table 1.5

Child Support Payments

Receipt State	Interviewer State: Between Waves SS	SP	PS	PP	Within Waves SS	PP
RR	.688 (1359)	.607 (82)	.547 (75)	.540 (67)	.733 (4647)	.651 (506)
RN	.056 (110)	.081 (11)	.073 (10)	.081 (10)	.024 (150)	.018 (14)
NR	.073 (145)	.096 (13)	.08 (11)	.081 (10)	.025 (160)	.022 (17)
NN	.183 (361)	.215 (29)	.299 (41)	.298 (37)	.218 (1379)	.309 (240)

Table 1.6

Private Pension

Receipt State	Interviewer State: Between Waves SS	SP	PS	PP	Within Waves SS	PP
RR	.874 (3304)	.878 (388)	.769 (286)	.819 (513)	.907 (11294)	.886 (2838)
RN	.014 (53)	.023 (10)	.008 (3)	.016 (10)	.004 (48)	.001 (4)
NR	.039 (147)	.036 (16)	.078 (29)	.045 (28)	.007 (84)	.004 (12)
NN	.073 (276)	.063 (58)	.145 (54)	.120 (75)	.083 (1030)	.109 (350)

Table 1.7

Supplemental Security Income

Receipt State	Interviewer State: Between Waves SS	SP	PS	PP	Within Waves SS	PP
RR	.813 (1293)	.745 (105)	.715 (98)	.831 (432)	.848 (439)	.850 (1685)
RN	.025 (40)	.057 (8)	.015 (2)	.015 (8)	.004 (22)	.004 (7)
NR	.036 (57)	.05 (7)	.08 (11)	.035 (18)	.01 (54)	.007 (13)
NN	.126 (200)	.149 (21)	.19 (26)	.119 (62)	.138 (714)	.14 (278)

Table 1.8

Unemployment Compensation and Benefits

Receipt State	Interviewer State: Between Waves SS	SP	PS	PP	Within Waves SS	PP
RR	.146 (485)	.151 (124)	.088 (63)	.104 (141)	.194 (2356)	.179 (1169)
RN	.123 (409)	.147 (120)	.128 (92)	.139 (189)	.049 (590)	.043 (282)
NR	.097 (324)	.111 (91)	.112 (80)	.082 (111)	.054 (660)	.041 (271)
NN	.634 (2107)	.591 (484)	.672 (481)	.675 (917)	.703 (8517)	.736 (4809)

APPENDIX 2

Effect of Imputation on Transitions

Table 2.1

Distributions of Transitions and Non-Transitions Between Waves

Source	Imputes Involved		Imputes Not Involved	
	Trans	Non-Trans	Trans	Non-Trans
Social Security	.039 (88)	.961 (2191)	.027 (499)	.973 (17785)
Veterans Compensation	.045 (11)	.955 (235)	.035 (67)	.965 (1854)
AFDC	.347 (43)	.653 (81)	.113 (257)	.887 (2014)
Food Stamps	.312 (74)	.688 (163)	.135 (773)	.865 (4940)
Child Support	.22 (24)	.78 (85)	.131 (296)	.869 (1966)
Private Pension	.103 (76)	.897 (663)	.049 (218)	.951 (4261)
Supplemental Security Income	.156 (23)	.844 (124)	.057 (128)	.943 (2113)
Unemployment Compensation	.459 (174)	.541 (205)	.192 (1119)	.808 (4720)

Trans = Transitions

Table 2.2

Distributions of Transitions and Non-Transitions Within Waves

Source	Imputes Involved: Trans	Imputes Involved: Non-Trans	Imputes Not Involved: Trans	Imputes Not Involved: Non-Trans
Social Security	.0003 (3)	.9997 (8594)	.008 (590)	.992 (75660)
Veterans Compensation	0 (0)	1.0 (711)	.004 (34)	.996 (8009)
AFDC	.003 (1)	.997 (301)	.030 (286)	.970 (9138)
Food Stamps	.007 (4)	.993 (596)	.043 (997)	.957 (22214)
Child Support	0 (0)	1.0 (335)	.052 (480)	.948 (8776)
Private Pension	0 (0)	1.0 (2130)	.01 (182)	.99 (18694)
Supplemental Security Income	0 (0)	1.0 (326)	.014 (125)	.986 (9121)
Unemployment Compensation	.232 (212)	.768 (701)	.108 (2616)	.892 (21656)

APPENDIX 3

Changes in Transition Distributions Due to Longitudinal Edit

Figure 3.1

Social Security

BETWEEN WAVE RECEIPT STATE TOTALS: UNEDITED

	MONTH 5	9	13
RR	6074	6243	6345
RN	92	71	66
NR	146	133	126
NN	484	349	259

WITHIN WAVE RECEIPT STATE TOTALS: UNEDITED

	MONTH 2	3	4	6	7	8	10	11	12	14	15	16
RR	6039	6077	6111	6211	6239	6279	6357	6371	6380	6458	6488	6514
RN	21	18	14	9	16	13	19	20	18	13	11	7
NR	56	48	55	44	53	35	34	27	31	41	33	32
NN	680	653	616	532	488	469	386	378	367	284	264	243

DIFFERENCE FROM UNEDITED TO EDITED
BETWEEN WAVE TOTALS

	MONTH 5	9	13
RR	6	-6	2
RN	-15	-10	-14
NR	-20	-1	6
NN	29	17	6

DIFFERENCE FROM UNEDITED TO EDITED
WITHIN WAVE TOTALS

	MONTH 2	3	4	6	7	8	10	11	12	14	15	16
RR	-3	6	4	-13	-14	-15	-7	-9	-10	8	7	4
RN	-8	-11	-8	-1	-4	-2	0	-1	0	0	0	0
NR	-2	-10	-13	-5	-3	-1	-3	-1	-2	-1	-3	-4
NN	13	15	17	19	21	18	10	11	12	-7	-4	0

Figure 3.2

Veterans Compensation

BETWEEN WAVE RECEIPT STATE TOTALS: UNEDITED

	MONTH 5	9	13
RR	640	653	655
RN	18	12	8
NR	26	11	11
NN	28	36	38

WITHIN WAVE RECEIPT STATE TOTALS: UNEDITED

	MONTH 2	3	4	6	7	8	10	11	12	14	15	16
RR	654	653	655	664	663	661	663	664	663	664	664	664
RN	0	4	0	2	3	3	1	0	2	2	1	1
NR	3	2	3	2	1	4	1	1	0	1	1	4
NN	55	53	54	44	45	44	47	47	47	45	46	43

DIFFERENCE FROM UNEDITED TO EDITED BETWEEN WAVE TOTALS

	MONTH 5	9	13
RR	3	-6	2
RN	-6	7	-7
NR	-1	1	-2
NN	4	-2	7

DIFFERENCE FROM UNEDITED TO EDITED WITHIN WAVE TOTALS

	MONTH 2	3	4	6	7	8	10	11	12	14	15	16
RR	1	1	0	2	2	2	-5	-5	-5	0	0	0
RN	0	-2	0	0	0	0	0	0	0	0	0	0
NR	-2	-1	-3	0	0	-1	0	0	0	0	0	0
NN	1	2	3	-2	-2	-1	5	5	5	0	0	0

Figure 3.3

AFDC

BETWEEN WAVE RECEIPT STATE TOTALS: UNEDITED

	MONTH 5	9	13
RR	517	531	531
RN	44	53	61
NR	61	57	44
NN	176	157	162

WITHIN WAVE RECEIPT STATE TOTALS: UNEDITED

	MONTH 2	3	4	6	7	8	10	11	12	14	15	16
RR	541	542	547	572	577	576	576	580	577	561	554	554
RN	9	14	12	6	10	13	12	11	11	14	16	17
NR	15	17	14	15	12	8	15	8	15	9	17	11
NN	233	225	225	205	199	201	195	199	195	214	211	216

DIFFERENCE FROM UNEDITED TO EDITED
BETWEEN WAVE TOTALS

	MONTH 5	9	13
RR	5	-6	2
RN	0	3	-6
NR	-9	1	-1
NN	4	2	5

DIFFERENCE FROM UNEDITED TO EDITED
WITHIN WAVE TOTALS

	MONTH 2	3	4	6	7	8	10	11	12	14	15	16
RR	2	5	7	-4	-3	-3	-4	-4	-3	2	3	3
RN	-1	-4	-3	0	-1	0	-1	0	-1	-1	-1	0
NR	-1	-1	-2	0	0	0	0	0	-1	0	0	-1
NN	0	0	-2	4	4	3	5	4	5	-1	-2	-2

Figure 3.4

Food Stamps

BETWEEN WAVE RECEIPT STATE TOTALS: UNEDITED

	MONTH 5	9	13
RR	1152	1160	1126
RN	192	167	168
NR	134	125	96
NN	474	500	562

WITHIN WAVE RECEIPT STATE TOTALS: UNEDITED

	MONTH 2	3	4	6	7	8	10	11	12	14	15	16
RR	1247	1256	1278	1259	1266	1275	1252	1251	1253	1195	1183	1194
RN	47	54	43	27	38	42	33	39	30	27	46	43
NR	63	65	66	45	51	52	38	32	41	34	54	43
NN	595	577	565	621	597	583	629	630	628	696	669	672

DIFFERENCE FROM UNEDITED TO EDITED BETWEEN WAVE TOTALS

	MONTH 5	9	13
RR	1	-6	2
RN	-20	-1	-11
NR	-9	-4	1
NN	28	11	8

DIFFERENCE FROM UNEDITED TO EDITED WITHIN WAVE TOTALS

	MONTH 2	3	4	6	7	8	10	11	12	14	15	16
RR	-17	-10	-13	-6	-5	-5	-9	-7	-7	5	6	9
RN	-8	-8	-9	-2	-2	0	-1	-3	-1	-2	-2	-3
NR	-1	-12	-6	-1	0	-2	-1	-1	-2	-1	0	-2
NN	26	30	28	9	7	7	11	11	10	-2	-4	-4

Figure 3.5

Child Support Payments

BETWEEN WAVE RECEIPT STATE TOTALS: UNEDITED

	MONTH 5	9	13
RR	508	520	521
RN	52	44	48
NR	56	52	81
NN	163	163	129

WITHIN WAVE RECEIPT STATE TOTALS: UNEDITED

	MONTH 2	3	4	6	7	8	10	11	12	14	15	16
RR	516	526	527	553	549	546	558	551	545	588	589	595
RN	19	22	32	11	19	18	14	30	24	14	22	17
NR	32	33	33	15	15	18	23	18	24	23	23	24
NN	212	198	187	200	196	197	184	180	186	154	145	143

DIFFERENCE FROM UNEDITED TO EDITED
BETWEEN WAVE TOTALS

	MONTH 5	9	13
RR	12	1	4
RN	-3	0	-2
NR	-12	0	-1
NN	3	-1	-1

DIFFERENCE FROM UNEDITED TO EDITED
WITHIN WAVE TOTALS

	MONTH 2	3	4	6	7	8	10	11	12	14	15	16
RR	4	5	12	1	2	2	1	2	2	4	4	4
RN	-6	-3	-13	-1	-1	0	0	-2	0	-1	-1	0
NR	-2	-6	-3	0	0	-1	-1	0	0	-1	0	-3
NN	4	4	4	0	-1	-1	0	0	-2	-2	-3	-1

Figure 3.6

Private Pension

BETWEEN WAVE RECEIPT STATE TOTALS: UNEDITED

	MONTH 5	9	13
RR	1398	1491	1551
RN	26	31	27
NR	107	68	64
NN	188	129	77

WITHIN WAVE RECEIPT STATE TOTALS: UNEDITED

	MONTH 2	3	4	6	7	8	10	11	12	14	15	16
RR	1409	1411	1408	1497	1503	1507	1553	1561	1568	1610	1615	1616
RN	9	6	9	8	5	11	6	8	6	5	1	8
NR	8	6	16	11	15	15	16	13	10	6	9	10
NN	293	296	286	203	196	186	144	137	135	98	94	85

DIFFERENCE FROM UNEDITED TO EDITED BETWEEN WAVE TOTALS

	MONTH 5	9	13
RR	15	1	2
RN	0	-6	-7
NR	-19	-4	-1
NN	4	9	6

DIFFERENCE FROM UNEDITED TO EDITED WITHIN WAVE TOTALS

	MONTH 2	3	4	6	7	8	10	11	12	14	15	16
RR	8	13	20	-3	-2	-2	-2	-3	-3	1	1	1
RN	-2	-5	-8	-1	-1	0	-1	-1	-2	0	0	0
NR	0	-1	-5	0	0	-3	-2	-2	-2	0	0	-1
NN	-6	-7	-7	4	3	5	5	6	7	-1	-1	0

Figure 3.7

Supplemental Securtiy Income

BETWEEN WAVE RECEIPT STATE TOTALS: UNEDITED

	MONTH 5	9	13
RR	601	641	669
RN	27	13	21
NR	36	37	34
NN	122	95	62

WITHIN WAVE RECEIPT STATE TOTALS: UNEDITED

	MONTH 2	3	4	6	7	8	10	11	12	14	15	16
RR	616	620	621	635	640	646	677	681	684	697	699	705
RN	7	4	7	2	3	4	1	1	5	6	5	4
NR	8	8	7	8	10	8	5	8	6	7	10	5
NN	155	154	151	141	133	128	103	96	91	76	72	72

DIFFERENCE FROM UNEDITED TO EDITED
BETWEEN WAVE TOTALS

	MONTH 5	9	13
RR	10	-1	2
RN	-3	0	-3
NR	-11	-1	-2
NN	4	2	3

DIFFERENCE FROM UNEDITED TO EDITED
WITHIN WAVE TOTALS

	MONTH 2	3	4	6	7	8	10	11	12	14	15	16
RR	5	6	8	-1	-1	-1	-2	-2	-1	0	1	1
RN	-4	-1	-4	0	0	0	0	0	-1	0	-1	0
NR	0	-2	-1	0	0	0	0	0	0	0	0	0
NN	-1	-3	-3	1	1	1	2	2	2	0	0	-1

Figure 3.8

Unemployment Compensation and Benefits

BETWEEN WAVE RECEIPT STATE TOTALS: UNEDITED

	MONTH 5	9	13
RR	292	272	211
RN	252	295	188
NR	236	157	166
NN	1262	1318	1477

WITHIN WAVE RECEIPT STATE TOTALS: UNEDITED

	MONTH 2	3	4	6	7	8	10	11	12	14	15	16
RR	454	417	402	428	457	445	344	317	302	282	281	306
RN	165	150	160	100	130	148	85	99	115	95	93	83
NR	113	145	142	159	136	122	72	100	97	92	108	114
NN	1310	1330	1338	1355	1319	1327	1541	1526	1528	1573	1560	1539

DIFFERENCE FROM UNEDITED TO EDITED BETWEEN WAVE TOTALS

	MONTH 5	9	13
RR	18	4	5
RN	20	10	28
NR	-4	25	14
NN	-34	-39	-47

DIFFERENCE FROM UNEDITED TO EDITED WITHIN WAVE TOTALS

	MONTH 2	3	4	6	7	8	10	11	12	14	15	16
RR	33	44	55	23	30	23	41	49	44	31	31	32
RN	-18	-14	-22	-9	-13	-2	-12	-14	-11	-12	-6	-6
NR	-3	-11	-17	-6	-9	-9	-6	-16	-11	-6	-5	-11
NN	-12	-19	-16	-8	-8	-12	-23	-19	-22	-13	-20	-15

APPENDIX 4

Investigation of Gross Changes in Income Recipiency from the Survey of Income and Program Participation

INVESTIGATION OF GROSS CHANGES IN INCOME RECIPIENCY FROM THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

Lynn Weidman, U.S. Bureau of the Census

INTRODUCTION

The Survey of Income and Program Participation (SIPP) is a longitudinal survey of households that collects economic information about the U.S. population. For two and one-half years the members of a household are interviewed at four month intervals and information is obtained for each of the four months preceding an interview. (This four month period is also called a "wave.") One type of estimate that can be derived from this monthly data is that of the number of people who change their response to a question between consecutive months or between any two fixed time points. A previous study (Burkhead and Coder, 1985) examined month-to-month changes in receipt of five different income types and two noncash benefits. It showed that, for the first twelve months of SIPP, the number of reported changes in recipiency status between the last month of one interview period and the first month of the next interview period was far greater than the number reported between any two months of the same interview period. Burkhead and Coder discussed these differences in relationship to questionnaire wording/design and respondent recall error.

In this investigation we are looking for more direct causes of the discrepancy in the between/within interview numbers of gross changes. (A gross change between two times is the number of people in state A at the first time and state B at the second time. The distributions of gross changes refers to these numbers for a specified set of pairs of states. We will be looking at reported gross changes only.) There are three phases of this investigation.

1. Empirical analysis of data to determine if demographic characteristics of individuals are related to the discrepancy.
2. Description and estimation of models for the effect of time in sample, recall lag and other sources of response error on reported gross changes.
3. Estimation of response error from outside sources and use of it in conjunction with the models.

Here we will present an empirical analysis and examine any significant results. Two models for relating error sources to gross changes are then proposed and presented for use in the next phase of investigation.

EMPIRICAL ANALYSIS

The goal of empirical analysis is to use simple methods to detect the existence of obvious relationships between demographic/interview characteristics and changes in receipt status of seven income types and food stamps. There are four receipt states for two consecutive months: RR, RN, NR and NN, where R = receipt and N = nonreceipt. The income types of interest are social security, unemployment compensation, private pensions, VA compensations and pensions, supplemental security income, child support and AFDC. They will be examined with respect to age, sex, race, marital status, education, relationship to principal person, household size, tenure, SMSA size and interview status. The distribution of gross changes in receipt status between consecutive months for each income type will be computed with respect to all pairs of demographic characteristics. This will produce 360 sets of distributions for examination. Any apparent relationships may suggest other distributions for examination.

The categories used for demographic variables are defined as follows.

age: 15-30, 31-45, 46-60, 61+
sex: male, female
race: white, nonwhite
education: elementary, high school, above high school
marital status: married, (separated, divorced, widowed), never married

household size: 1,2,3,4-5,6+
tenure: home owned, not owned
relationship to reference person: reference person, spouse, child, other
SMSA size: not in an SMSA, 1 million +, less than 1 million
interview status for consecutive months: SS,SP,PS,PP where S=self, P=proxy

The file of monthly data was created from the first four waves of data available for each household. Each of these waves is searched for all persons who reported receipt of any of the income types of interest during any month of the wave. For each such person all the information available for the 16 month period is collected and placed on a record. This record will then be used if the person was interviewed for each of the four waves. (Restricting the analysis to these persons follows the Burkhead and Coder data set selection for the first twelve months.) A wave on the record was then used only if it was preceded by a wave of matching data. This ensures that the last three months of a wave are used in the calculations only if the first month is also. (An important fact to remember is that the large majority of people are not included on this file because they do not receive any of these income types.)

How will we determine if any relationships exist? When the monthly gross changes are computed there are usually two to five times as many RN and NR reported for the first month of a wave as there are for the other three months. (See Table 1.) For any pair of demographic variables to be a determinant of this change, we would have to observe a huge difference in the number of RN and NR reported in the first months of waves as compared to the last three months for some combination(s) of these variables, but not for others. We will be looking for one or more combinations to exhibit this behavior.

As a theoretical example of the distributions that were calculated see Table 2. There are two such tables for each comparison. The first is for all first months of a wave combined (between waves) and the second is for all months two, three, and four combined (within waves). This means that the total number of observations in the second table is three times the number of observations in the first.

TABLE 2
RACE

		white		non-white	
SEX	male	P_1RR P_1NR	P_1RN P_1NN	P_2RR P_2NR	P_2RN P_2NN
	female	P_3RR P_3NR	P_3RN P_3NN	P_4RR P_4NR	P_4RN P_4NN

Within each cell defined by a particular combination of demographic characteristics we calculate the probability of each receipt state, $P_1AB=P$ (receipt state AB/cell 1). Let P_1AB_w denote such a probability within waves and P_1AB_b the corresponding between wave probability. Compare P_1NR and P_1RN for between waves to those for within wave. If this demographic combination has no relationship to gross changes, the ratios P_1NR_b/P_1NR_w should be fairly constant for 1, as should the ratios P_1RN_b/P_1RN_w. If one and/or both of these sets of ratios differ "greatly" between cells, this indicates the type of relationship we are looking for. (It is important to note that no statistical tests were performed. Comparisons are made by examining distributions for specified types of "noticeable" differences.)

When examining interview status the situation is somewhat different because two of the interview status pairs, PS and SP, cannot occur within waves. In this case we look for large differences in the distributions of P_1NR_b and P_1RN_b between cells.

Examination of these tables showed no major relationships between demographic variables and the gross changes. Some small differences in distributions occur, but nothing on the order of magnitude of the between/within wave gross change differences. As an example, see Table 3, sex x race for food stamps.

TABLE 3.A

Food Stamps: Between Waves
Race x Sex

Race	Sex	RR	RN	NR	NN
white	male	44.3 (547)	11.8 (146)	6.1 (75)	37.9 (468)
	female	59.7 (1560)	7.8 (205)	6.2 (163)	26.2 (684)
non-white	male	54.0 (262)	10.3 (50)	7.6 (37)	28.0 (136)
	female	68.9 (1086)	6.2 (97)	4.7 (74)	20.3 (320)

TABLE 3.B

Food Stamps: Within Waves
Race x Sex

Race	Sex	RR	RN	NR	NN
white	male	49.3 (1830)	2.0 (73)	3.1 (116)	45.6 (1695)
	female	64.2 (5031)	2.0 (154)	2.2 (172)	31.6 (2479)
non-white	male	61.2 (891)	1.4 (20)	1.6 (23)	35.8 (521)
	female	72.6 (3433)	1.4 (64)	1.7 (79)	24.4 (1155)

First entry in each cell is percent of total responses in row. Second entry is number of responses in cell.

Food stamps, social security and unemployment compensation were the sources with relatively large numbers of transitions reported. (I.e., with enough transitions to compare distributions for many cells.) The first two of these sources showed about the same patterns. Larger proportions of receipt of sources were reported by self-respondents than by proxies. There is usually a higher proportion of transitions between waves when at least one of two consecutive months has a proxy response than when both of the months are self-reported. As an example, see Table 4. Because the number of SS cases was much larger than the sum of SP, PS, and PP cases, these patterns did not have a noticeable effect on the within/between wave jumps. (For unemployment compensation there is a much larger number of cases with NN. The patterns are similar, but the difference in proportions are much smaller.)

TABLE 4.A

Food Stamps: Between Waves
Sex x Interview State

Sex	Interview State	RR	RN	NR	NN
Male	SS	54.5 (456)	9.4 (79)	6.0 (50)	30.1 (252)
	SP	45.7 (106)	12.5 (29)	8.6 (20)	33.2 (77)
	PS	38.2 (76)	16.1 (32)	8.0 (16)	37.7 (75)
	PP	37.7 (171)	12.4 (56)	5.7 (26)	44.2 (200)
Female	SS	65.5 (2326)	6.8 (240)	5.2 (184)	22.6 (802)
	SP	53.9 (125)	9.1 (21)	8.5 (20)	28.4 (66)
	PS	43.1 (103)	9.2 (22)	9.2 (22)	38.4 (92)
	PP	55.4 (92)	11.4 (19)	6.6 (11)	26.5 (44)

TABLE 4.B

Food Stamps: Within Waves
Sex x Interview State

Sex	Interview State	RR	RN	NR	NN
Male	SS	57.3 (1782)	1.5 (47)	2.5 (77)	38.7 (1202)
	PP	45.7 (939)	2.2 (46)	2.7 (56)	49.3 (1014)
Female	SS	68.1 (7750)	1.7 (198)	2.1 (236)	28.0 (3189)
	PP	59.8 (714)	1.7 (20)	1.3 (15)	37.3 (445)

MODELS

Since the empirical analysis failed to reveal any relationships between demographic variables and the distribution of gross changes, we must look for another way of determining their true distributions. For CPS it has long been known that there is a relationship between the responses to a question and (1) the amount of

time that has elapsed between the month of interest and the month of interview, (ii) the interview status and (iii) the length of time a person has been in the sample. Here we propose models for gross changes that make use of similar relationships.

The dependent variable of interest for a given income type is the receipt state identified with the second of two consecutive months. The possible receipt states for month t are (1)=RR, (2)=RN, (3)=NR, (4)=NN. Let $y_{ijkt(\ell)}$ be the number of responses in receipt state ℓ in month t where

- i = number of times a person has been interviewed,
- j = number of months between month t and month of interview,
- k = interview status for months t-1 and t; PP,PS,SP and SS with S=self, P=proxy.

Then the vector $\underline{y}_{ijkt}$ = $(y_{ijkt(1)}, y_{ijkt(2)}, y_{ijkt(3)}, y_{ijkt(4)})$ represents the gross change counts for the combination ijkt.

Multivariate Normal Models

Since the $\underline{y}_{ijkt}$ are vectors of counts, they have a multinomial rather than a multivariate normal distribution. But because of the large sample sizes on which they are based (the total number of counts in y_{ijkt}), they have that distribution asymptotically. We propose a multivariate analysis of variance (MANOVA) model of the form

$$E(y_{ijkt(\ell)}) = u_{(\ell)} + M_{i(\ell)} + M_{j(\ell)} + S_{k(\ell)} + MM_{ij(\ell)} + MS_{ik(\ell)} + MS_{jk(\ell)} + \tau_t \qquad (1)$$

where the terms are

- M_i = interview number i,
- M_j = months of recall between month of interview and month of occurrence,
- S_k = interview status,
- MM_{ij}, MS_{ik}, MS_{jk} are interactions of these effects, and
- τ_t = month t.

There are some difficulties we must take account of before using this model.

(1) Levels 2 and 3 of k occur only with j=4. This means that the cells which are defined with j=4 and k=1 or 4 contain structural zeros. The contrasts in the analysis that define the effects and their degrees of freedom must be consistent with these structural zeros.

(2) The effect for interview number is to determine if reporting of changes in state follows some pattern over time. For example, a person may report the specific month of transition in wave 1, but after that he reports all transitions as occuring in the first month of a wave. Suppose now that there is a proxy respondent for waves 2 and 3. Will the proxy behave as the self respondent did for wave 1, or as he would for wave 2, or in some different manner? In a strict sense this effect only has validity if the same respondent is available in each wave. However, we can still include this effect as an average response difference between successive interviews.

(3) Most of the data that is used in this modeling is not available on the file we are using. Recall that only persons who have received one of the eight income sources in the first 16 months of SIPPS are included in this file. The vast majority of persons have no receipt for the first 16 months and would thus have the receipt state NN for each of the months used in modeling. From the files for individual waves we would have to calculate the number of these persons in each cell defined by an ijkt combination. The most time-consuming part of this job would be matching records across waves.

Polytomous Logit Models

There is another approach we can take to this problem that does not require a multivariate normal distribution. Instead of modeling the frequency of each receipt state we can model the probabilities of the states with polytomous logit models. A brief description of these models is given.

Let an observation consist of a set of

independent variables $\underline{x}_i$ and a dependent variable y_i, where y_i falls into one of 6 mutually exclusive categories. Let $\underline{\beta}_g$ be a set of coefficients for category g, g=1,2,...6. Assume that

$$\text{Prob}\,(y_i=g) = \exp(\underline{x}_i'\underline{\beta}_g)/\sum_{g=1}^{6}\exp(\underline{x}_i'\underline{\beta}_g). \qquad (2)$$

The unknown $\underline{\beta}_g$, g=1,2,...6, can be estimated by maximum likelihood, where the likelihood function is

$$[\prod_{i=1}^{N}\exp(\underline{x}_i'\underline{\beta}_{h(i)})]/[\sum_{g=1}^{6}\exp(\underline{x}_i'\underline{\beta}_g)]^N$$

and h(i) is the category into which y_i falls. Note that the probability in (2) remains constant if all $\underline{\beta}_g$ are multiplied by a constant, so a single linear restriction must be placed on the $\underline{\beta}_g$'s to obtain unique maximum likelihood estimates.

We propose using this logit model approach to estimate the true proportion of responses in each receipt state at each time t. Let $\underline{x}_{ijkt}$ be the vector of 0-1 variables that indicate which main effects and interactions are present for each observation with a particular ijkt combination. Let $\underline{\beta}_s$ be the vector of corresponding effects for receipt state s. Each observation that is counted in $y_{ijkt(s')}$ will contribute a term of the form

$$\exp(\underline{x}_{ijkt}'\underline{\beta}_s)/\sum_{s=1}^{4}\exp(\underline{x}_{ijkt}'\underline{\beta}_s) \qquad (3)$$

to the likelihood function. Thus we only need to compute all the $\underline{x}_{ijkt}$ in order to determine the likelihood function and the resulting maximum likelihood estimates $\hat{\underline{\beta}}_s$, s=1,2,3 or 4. Then the estimated proportion of observations in receipt state s for combination ijkt is obtained by substituting the $\hat{\underline{\beta}}_s$ into (3).

The same difficulties that were described for MANOVA models are also present here.

When using either of these modeling approaches we would test for main effects and interactions being zero in order to determine which of them influence the reporting of changes in receipt. For MANOVA models standard procedures are available and for logit models likelihood ratio tests are used for nested models; i.e., for testing that certain entries in $\underline{\beta}_s$, s=1,2,3,4, are zero.

SUMMARY

An empirical examination did not detect any relationships between gross change distributions and nine demographic variables and interview status. Modeling approaches are proposed for estimating the true number and proportion of each receipt state for a particular combination of interview number, months recall, interview status and month. Tests of significance for main effects and interactions can be carried out to determine which of them influence reporting of changes in receipt status. The resulting models could be used to adjust the reported gross changes toward the actual gross changes. More consideration of the validity of the models and the amount of work required to carry out estimation needs to be done before carrying this work further.

Mention should be made of another study that is in progress at the Census Bureau. A comparison of administrative records obtained from four states with SIPP data is being made to investigate the relationship between reported and actual changes in status. We hope to be able to use these results in conjunction with models to get an improved estimate of gross change distributions.

REFERENCES

Burkhead, Dan and John Coder (1985). "Gross Changes In Income Recipiency From the Survey of Income and Program Participation." American Statistical Association, Proceedings of the Social Statistics Section.

Hogue, Carma (1984). "History of the Problems Encountered in Estimating Gross Flows." Proceedings of the Conference on Gross Flows in Labor Force Statistics, Washington, D.C., 1-5.

Merlove, Marc and S. James Press (1973). Univariate and Multivariate Log-linear and Logistic Models. Rand Corporation, R-1306-EDA/NIH.

TABLE 1

Month-to-Month Gross Changes: Food Stamps

Receipt Status	1st to 2nd	2nd to 3rd	3rd to 4th	4th to 5th	5th to 6th	6th to 7th	7th to 8th	8th to 9th	9th to 10th	10th to 11th	11th to 12th	12th to 13th	13th to 14th	14th to 15th	15th to 16th
RR	1240	1255	1274	1159	1270	1278	1287	1161	1260	1261	1265	1135	1216	1205	1219
RN	40	47	35	174	26	38	42	167	33	36	29	157	25	44	40
NR	62	54	61	129	46	51	51	123	37	33	40	97	33	54	43
NN	653	639	625	517	652	627	614	519	659	659	655	572	713	684	685