

RESPONSE ERRORS IN SIPP:
SOME BEGINNING ESTIMATES BASED ON ADMINISTRATIVE RECORDS

Jeffrey C. Moore and Kent H. Marquis
U.S. Bureau of the Census, Washington, D.C. 20233

KEY WORDS: response bias, record check

I. INTRODUCTION

We want to know how well the Survey of Income and Program Participation (SIPP) measures the personal economic characteristics and changes in characteristics of U.S. residents. The quality of such measures can have important effects on the conclusions that data users draw. Quantitative knowledge about data quality can assist users in correcting for the errors, and can help survey designers construct improved measurement procedures.

To address the measurement quality issues, the Census Bureau is conducting a record check study in four states, using records from nine different programs and matching them to sample persons in the first two waves of the 1984 panel of SIPP.

In this paper we present two kinds of bias estimates: the discrepancy score and the seam bias. The discrepancy score is the average difference between the survey value and the true value from administrative records. The seam bias refers to the tendency for change rates to be larger for time periods between interviews than for time periods measured in the same interview. The seam bias is a new evaluation concept for longitudinal surveys. We estimate it here and use the administrative record data to attempt some beginning understanding of it.

We provide the response bias estimates for four subject matter characteristics (program participation rates, average program benefit amounts, participation transition rates, and benefit amount change rates) for two programs (AFDC and Food Stamps) in one state (Wisconsin).

In section II we give a brief description of SIPP, the record check study design, the administrative records and the record linkage procedures. In section III we discuss the subject matter characteristics of interest and how we estimate the response biases. In section IV we present and discuss the response bias estimates, indicating that one of the programs is well-reported and the other contains several kinds of response bias. And in the last section we discuss the implications of the results for both SIPP and our future error modeling work.

II. BACKGROUND

The purpose of SIPP is to provide improved information on the economic situation of persons and families in the United States by collecting comprehensive longitudinal data on a wide range of topics including cash income, noncash income, eligibility for (and participation in) government transfer programs and family dynamics. Each sample household is interviewed by personal visit eight times--once every four months for two-and-one-half years. At each visit to the household, each person 15 years of age or older is asked to provide information about himself/herself, but proxy reporting is permitted. The reference period for each interview is the four months preceding the interview month. In this

paper we limit analysis to the first two interviews (eight months) of the 1984 longitudinal panel.

SIPP uses a "rotating" panel design. Each panel is divided into four groups and the groups are interviewed in successive months. For the 1984 panel, the reporting period began in October 1983 for the first rotation group, in November 1983 for the second group, etc. In our analyses we group results according to SIPP reference month--e.g., wave 1 month 4 (four months ago), wave 1 month 3 (three months ago), etc. Readers should keep in mind that the rotation group interviewing structure means that each reference month includes data from four calendar months.

Elsewhere (Moore and Marquis, 1987) we have discussed the major issues to be faced in designing a record check evaluation of survey measures. A key characteristic of this record check is its "full design," in which we obtain records for all people in the survey (or state), not just those who claim to have the recorded characteristic in the interview. Similarly, we interview a sample of households in the state and the sample is drawn independently of values in the administrative records.

The project uses records from five Federally administered programs (Civil Service Retirement, Pell (education) Grants, Social Security, Supplemental Security Income, and Veterans Compensation/Pensions) and four state-administered programs (Aid to Families with Dependent Children, Food Stamps, Unemployment Compensation, and Workers' Compensation). Confidentiality provisions in the United States Code permit the Census Bureau to receive such data from other agencies for research purposes, but prohibit the disclosure of any information about individuals.

The study uses computerized matching procedures based on Fellegi and Sunter's (1969) theoretical work on record linkage. First, matching information is put into standard formats. Then the survey and administrative files are sorted into small subsets of records (blocks). Step three involves estimating two match parameters for each match variable, and the final step is to implement the actual record matching on the computer.

III. CONSIDERATIONS IN MAKING THE ESTIMATES

In this paper we estimate response bias for several characteristics for two programs administered by the state of Wisconsin--Aid to Families with Dependent Children (AFDC) and Food Stamps. In this section we describe these programs and some of the major considerations in estimating the various bias parameters.

PROGRAMS

Aid to Families with Dependent Children (AFDC)

AFDC is an income-tested program providing money to needy families with children, including single-parent families and 2-parent families where one parent is disabled or the father is

unemployed. We consider the household to be participating in AFDC if anyone in the household reports participation. Household benefit amounts are formed by combining the amounts reported by all household members, except for administrative record cases in which duplicate amount entries are clearly flagged. (Instances of multiple amount reporting in SIPP are very rare.)

Food Stamps

The Food Stamps program provides coupons to low-income households to purchase a nutritionally adequate diet. Virtually all types of households are eligible, and the amount received is determined by a formula that includes income. Scoring rules for participation and amounts are the same as for AFDC.

UNIT OF ANALYSIS

The unit of analysis for these two programs is the longitudinal household, rather than the individual. For this analysis, a household remained the same from month to month so long as three conditions held: (1) it retained the same head (and, if present, same spouse); (2) the head or spouse was part of the household during the first interview month of the panel; and (3) it remained the same household type. There are 5 household types: (1) married couple household; (2) other family household with a male head; (3) other family household with a female head; (4) nonfamily household with a male head; and (5) nonfamily household with a female head.

SUBJECT MATTER CHARACTERISTICS

Participation Rate

The sample's participation rate for a time period is defined as p/n , where p is the number of households receiving the benefit during the time period and n is the number of households in the sample. P may be determined from either the survey or the administrative records.

Average Benefit Amount

The sample's average benefit amount is $\Sigma b / c$ where Σb is the sum of the reported or recorded benefit amounts for a subset, c , of the sample. C is the number of households with nonzero amounts in both the interview and administrative record. We make this restriction in order to estimate how much error occurs in benefit amount reporting apart from errors in participation reporting.

Participation Transition Rate

We infer a participation transition rate from participation data for two adjacent months. A transition between months occurs when there is a change in participation status between months. The sample's participation transition rate is t/n , where t is the number of households that changed between the adjacent months and n is the number of households present in both months.

Amount Change Rate

The amount change is just $\Sigma d / n$, where $d=0$ if the amount the household received in the first month is the same as the amount received in the second month, and $d=1$ if the amounts the household received in the first and second months differ. N is the number of households

participating in both months according to either the interview or the administrative record data. The Σ indicates summation over households.

RESPONSE BIASES

For each of the four characteristics and both of the programs we estimate the average response bias using the record data. For the two transition characteristics we also estimate a "seam bias" using only the survey data.

Average Response Bias

The average response bias is the average discrepancy between survey and record values for the households in the sample: $\Sigma_i (S_i - R_i) / N$, where S is the survey value, R is the administrative record value and i indexes the N households in the sample. (The definition assumes that the record values represent truth, an assumption we have no reason to question for these data.)

We estimate a confidence interval for the average response bias as $\pm [\text{VAR}(S_i - R_i) / N]^{1/2}$, the square root of the variance of the survey minus record discrepancies divided by the number of households in the sample. This estimator does not take into account the possible nonindependence of the sample observations due, for example, to cluster sampling. However, the estimator leads to meaningful conclusions when the uncorrected version includes zero because a corrected version would also encompass zero. Thus, in the analyses to follow, we make unqualified assertions about lack of "statistical significance" but attach qualifications to estimates that appear to be statistically significant.

To make interpretation of the numerical results easier, we occasionally refer to the percent bias, which is standardized by dividing the average response bias by the record (true) average for the characteristic. Thus, for example, if the true participation rate in the sample was 20% and the reported rate was 15%, the percent bias would be $(15\% - 20\%) / 20\% = -25\%$.

Seam Bias

Several researchers (Moore and Kasprzyk, 1984; Burkhead and Coder, 1985; Hill, 1987) have noted the tendency for more changes (or higher rates of change) to be observed between months covered by two interviews than between the months covered by any one interview. We refer to this as the seam bias. We do not actually estimate a seam bias; descriptively we just point to the different measured amounts or rates of change measured across versus within interviews. The inferential statistic is a paired comparison t -test that compares the nonseam transition rate (an average of six values) to the seam transition rate using the longitudinal households that participated and did not change during the eight months of the survey.

IV. RESULTS

Results of our initial analyses suggest that response errors operate quite differently for the two programs. In general, AFDC participation and amount reporting appear unbiased, on average. Food Stamps reports, on the other hand, exhibit error patterns which clearly indicate response bias.

AFDC

We discuss in detail below the general finding of relatively low levels of average response bias in reports of AFDC characteristics. First we examine bias in each of the monthly rates of participation and benefit amounts reported. These biases are not significantly different from zero. Then we look at the bias in the monthly changes inferred for participation and benefit amounts. While data trends are in the direction one would expect if there were a seam bias, only a few of the expected differences are statistically significant.

AFDC--Monthly Participation and Amounts

For each month covered by the two SIPP interviews, Table 1 indicates that the estimated average response bias for the participation rate is zero. Each bias estimate is within the range of $\pm 12\%$ of "truth" and all of the confidence intervals include zero. There is no consistent directional pattern to the rate discrepancies. The data suggest neither consistent underreporting nor consistent overreporting nor memory decay (decreasing reporting quality with increasing recall length).

Similarly, none of the estimated biases in reported AFDC benefit amounts is significantly different from zero, as shown in Table 2. The percent bias estimates are in the very narrow range of $\pm 4\%$. Again, there are no systematic error trends apparent such as consistent overreporting, underreporting, or memory decay.

Thus, reports of the monthly AFDC characteristics by Wisconsin households in the first eight months of SIPP appear unbiased on the average.

AFDC--Changes

The seam bias effect for AFDC participation (Table 3) is in the expected direction, but a t-test comparing the seam change rate (.0078) with the unweighted average nonseam rate (.0032) is not statistically significant (for $\alpha = .05$). In addition, only one of the seven mean discrepancies (survey rate minus record rate) in Table 3 is statistically significant, indicating that these SIPP reports of AFDC participation transitions are, in general and on average, unbiased.

The AFDC benefit amounts reported in SIPP, however, do show a seam effect (Table 4). At the seam, about 36% of SIPP households who reported receipt in both months reported an amount change, versus an average rate of only about 13% for nonseam month pairs. What do the records say about the seam bias--are the reports at the seam correct, as some have speculated, or are there too many changes inferred at the seam and too few reported elsewhere?

The record check data in Table 4 do not provide a very definitive answer. The trend suggests that SIPP elicits too many changes at the seam, but the t value doesn't quite reach the critical value of 1.96 (and would actually be smaller if we had taken the sample design effects into account in estimating the variances). Furthermore, while the trend is for all of the nonseam biases to be negative (too few transitions reported for nonseam months),

only one of the six bias estimates, based on record data, is statistically significant.

Thus, the reported changes in AFDC characteristics show a little more average reporting bias than the monthly point estimates, but in general the change reporting biases are not significantly different from zero either. We don't detect a conventional seam effect for changes in participation, but we do find one (almost) for changes in benefit amounts. There is little or no average bias in reports of participation changes, but some evidence that benefit amount change rate estimates for AFDC are biased.

FOOD STAMPS

In the discussion below, we show several kinds of reporting bias for Food Stamp program characteristics in SIPP (for Wisconsin households in the time period studied). We find a consistent tendency to underreport participation in the program, although if participation is reported correctly, so are the benefit amounts on the average. In addition, the participation data suggest that the forgetting biases may have a "decay" component, with the biases getting larger as the recall period lengthens.

Both participation and benefit amount changes show a seam bias effect in SIPP: more changes are inferred at the seam than for the other pairs of months. The administrative record data suggest (with some qualifications) that there are too many changes reported at the seam and too few in the other months.

FOOD STAMPS--Monthly Participation and Amounts

The average biases in Food Stamps participation rates appear, from Table 5, to be moderately large and negative. The percent biases range from 0 to about -26% and the confidence intervals for six of the eight biases do not include zero. In addition, it appears as if the bias generally increases as the elapsed time between the interview and the reference month increases, which we term a "memory decay" pattern. In wave one the percent bias trend goes from -12.5% in month 1 ("last month") to -18.4% in month 4 ("four months ago"). Similarly, for wave 2, the percent bias estimate for the shortest recall period (month 1) is 0 and the estimate for the longest recall (month 4) is -24.2% . We note that in each of the waves one of the estimates among the four months is a little out of line, and that we have not attempted to fit any particular model to the data in order to assess the statistical significance of the trend. But we are alerted to the possibility of a time-dependent form for a forgetting bias and will take that into account in future work.

Unlike the monthly participation reports, there seems to be little or no bias in reporting the monthly value of food stamps received, conditional on correct reporting of participation (Table 6). The bias estimates range from $+14\%$ to -9.7% for the eight months studied, but none of the estimates is significantly different from zero, and none of the trends that we have looked for in other tables is evident (consistent overreporting, consistent underreporting, or memory decay). So there is underreporting of Food Stamps participation, but if households reported

participation correctly then they reported the benefit amount correctly on the average.

FOOD STAMPS--Changes

There is a clear seam effect in reported Food Stamps program participation changes. The change rate at the seam is .0195 (Table 7), about three times the average rate (.0068) for nonseam periods. For this variable, however, the record information is not very informative about the underlying response errors: the trend is for too many changes reported at the seam, but the confidence interval for this estimate includes zero ($t = 1.94$) even without taking into account the effects of the sample design on the estimated variances. Four of the six bias estimates for the nonseam months are negative, but none is significantly different from zero. So, while the data suggest consistency with typical seam effects, nothing is statistically significant.

The seam bias and its underlying reporting biases are a little clearer for changes in Food Stamps amounts (Table 8). The seam change rate of 75% significantly exceeds the 29% average for nonseam time periods. Using the administrative record data, it appears that the true change rate at the seam was only 41%, resulting in a substantial net overreporting of change inferred for that time period. Five of the six bias estimates for the nonseam periods are negative and four of six are statistically significant (perhaps fewer would be statistically significant if we had paid attention to sample design effects). There are no instances of a positive net response bias in the nonseam periods.

So, these data seem to confirm what was suggested by the AFDC amount change results--when we have a seam bias in the survey, we are inferring too many changes between interviews and too few changes for the time periods within a single interview.

V. SUMMARY AND DISCUSSION

The results indicate that AFDC reports (in Wisconsin for the first eight months of SIPP) were fairly accurate on the average while there were potentially important response problems with the Food Stamps reports. These results are possible because we could compare the interview reports to the state AFDC and Food Stamps administrative records for all SIPP households in the state. These results, however, may not generalize to other states, programs and times.

Despite their limitations, these results have important implications for how we will address later descriptive modeling and hypothesis testing on the complete sample. For example, we now know that we cannot automatically combine data across programs; we will instead attempt to model separately the characteristics of programs that are reported well and poorly.

Learning how the basic response errors produce a seam effect will be a substantial challenge. With so few cases, we came close to detecting a seam bias only part of the time. The administrative record data suggest both a net overreporting of changes at the seam and a net underreporting of changes at other times. Modeling that pattern of biases will be difficult, and we suspect that our models will encompass more than the simple average bias parameter mentioned here. Our

measurement models may need several error parameters to describe both the monthly discrepancies (mean and variance) and the month-to-month reporting errors. Our current thinking is to model the error variance such that it is highly correlated among months within an interview (wave), and much less so between interviews.

Finally, we need to expand our treatment of errors in reporting benefit amounts and changes in amounts. The results presented here, conditional on correct reporting of participation, suggest that there are only small average errors in reporting of amounts, but that there may be substantial bias in the reporting of amount changes. This suggests that we may adopt a two-part model of response errors, the first part describing errors in participation reporting and the second describing errors in attributes, conditional on the participation reporting error.

REFERENCES

- BURKHEAD, D. and J. CODER (1985). "Gross Changes in Income Reciprocity from the Survey of Income and Program Participation," Proceedings of the Social Statistics Section, American Statistical Association, Washington, DC.
- FELLEGI, I. and A. SUNTER (1969). A Theory for Record Linkage. Journal of the American Statistical Association, Vol. 64, pp. 1183-1210.
- HILL, D. (1987). "Response Errors Around the Seam: Analysis of Change in a Panel with Overlapping Reference Periods," Proceedings of the Survey Research Methods Section, American Statistical Association, Washington, DC.
- MOORE, J. and D. KASPRZYK (1984). "Month-to-Month Reciprocity Turnover in the ISDP," Proceedings of the Survey Research Methods Section, American Statistical Association, Washington, DC.
- MOORE, J. and K. MARQUIS (1987). "Using Administrative Record Data to Evaluate the Quality of Survey Estimates." Paper presented at Statistics Canada's International Symposium on Statistical Uses of Administrative Data, Ottawa, Ontario, Canada, November 1987.

Acknowledgements

The research program described in this paper has benefited greatly from the contributions of many people. While we cannot list here all who deserve recognition, we do gratefully acknowledge the particular contributions of: Bill LaPlant, for his assistance regarding the operation of the Census Bureau's computer matcher; Chris Dyke, for his guidance on the use of the computer system used to run the matcher; Elaine Fansler and Jeannette Robinson, for their efforts in preparing the multitude of files for matching and for executing the match runs; Doris Anderson, for her assistance in preparing the many SIPP files required by the project; Don Hernandez, for his assistance on household (and longitudinal household) definitions; Dan Gillman, Betsy Martin, and Glen Wolfgang, for their comments on early drafts of this paper; and Dan Kasprzyk, for his constant support of the entire endeavor.

Table 1: AFDC Participation According to SIPP and Administrative Records

Wave-Month	(N)	Participation Rate:		Mean Discrepancy	Std. Error of Mean	t	Percent Bias
		SIPP	Record				
1-4 mos. ago	(529)	.0510	.0491	.0019	.0033	0.58	3.8%
1-3 " "	(529)	.0529	.0586	-.0057	.0050	-1.13	-9.7%
1-2 " "	(529)	.0510	.0548	-.0038	.0046	-0.82	-6.9%
1-1 " "	(529)	.0510	.0548	-.0038	.0046	-0.82	-6.9%
2-4 " "	(523)	.0458	.0516	-.0057	.0043	-1.34	-11.1%
2-3 " "	(530)	.0490	.0490	0	.0027	0	0
2-2 " "	(530)	.0509	.0509	0	.0027	0	0
2-1 " "	(534)	.0505	.0486	.0019	.0019	1.00	3.8%

Table 2: AFDC Benefit Amounts According to SIPP and Administrative Records for Households With Nonzero Amounts in Both SIPP and the Records

Wave-Month	(N)	Mean Benefit Amount:		Mean Discrepancy	Std. Error of Mean	t	Percent Bias
		SIPP	Record				
1-4 mos. ago	(25)	\$537.08	533.53	3.55	7.00	0.51	0.7%
1-3 " "	(26)	518.62	518.92	-0.31	5.19	-0.06	-0.1%
1-2 " "	(25)	531.40	519.93	11.47	11.42	1.00	2.2%
1-1 " "	(25)	516.24	507.69	8.55	9.86	0.87	1.7%
2-4 " "	(23)	570.13	556.54	13.59	12.94	1.05	2.4%
2-3 " "	(25)	545.00	548.74	-3.74	21.97	-0.17	-0.7%
2-2 " "	(26)	527.58	508.16	19.41	15.57	1.25	3.8%
2-1 " "	(26)	530.73	524.43	6.30	13.23	0.48	1.2%

Table 3: Month-to-Month AFDC Participation Transitions According to SIPP and Administrative Records

Wave-Month Pair:	Transition Rate:		N	Mean Discrepancy	Std. Error of Mean	t	
	SIPP	Record					
1:	1-4 to 1-3	.0019	.0132	529	-.0113	.0053	-2.13
	1-3 to 1-2	.0057	.0076	529	-.0019	.0050	-0.38
	1-2 to 1-1	0	.0038	529	-.0038	.0027	-1.41
Seam:	1-1 to 2-4	.0078	.0078	513	0	.0055	0
2:	2-4 to 2-3	.0039	.0058	514	-.0019	.0043	-0.44
	2-3 to 2-2	.0057	.0019	523	.0038	.0038	1.00
	2-2 to 2-1	.0019	.0039	518	-.0019	.0033	-0.58

Table 4: Month-to-Month AFDC Benefit Amount Transitions According to SIPP and Administrative Records for Households With Nonzero Amounts in Both Months

Wave-Month Pair:	Transition Rate:		Mean Discrepancy	Std. Error of Mean	t	
	SIPP (N)	Record (N)				
1:	1-4 to 1-3	.0741 (27)	.3600 (25)	-.2859	.1144	-2.50
	1-3 to 1-2	.1538 (26)	.2143 (28)	-.0604	.1055	-0.57
	1-2 to 1-1	.2222 (27)	.2500 (28)	-.0278	.1145	-0.24
Seam:	1-1 to 2-4	.3636 (22)	.1250 (24)	.2386	.1268	1.88
2:	2-4 to 2-3	.1250 (24)	.2500 (24)	-.1250	.1127	-1.11
	2-3 to 2-2	.1600 (25)	.1923 (26)	-.0323	.1067	-0.30
	2-2 to 2-1	.0385 (26)	.1200 (25)	-.0815	.0757	-1.08

Table 5: Food Stamps Participation According to SIPP and Administrative Records

Wave-Month	(N)	Participation Rate:		Mean Discrepancy	Std. Error of Mean	t	Percent Bias
		SIPP	Record				
1-4 mos. ago	(529)	.0586	.0718	-.0132	.0056	-2.34	-18.4%
1-3 " "	(529)	.0604	.0699	-.0095	.0042	-2.24	-13.5%
1-2 " "	(529)	.0548	.0623	-.0076	.0038	-2.01	-12.1%
1-1 " "	(529)	.0529	.0604	-.0076	.0038	-2.01	-12.5%
2-4 " "	(523)	.0478	.0630	-.0153	.0060	-2.54	-24.2%
2-3 " "	(530)	.0433	.0584	-.0151	.0065	-2.32	-25.8%
2-2 " "	(530)	.0452	.0509	-.0057	.0042	-1.34	-11.1%
2-1 " "	(534)	.0505	.0505	0	.0037	0	0

Table 6: Food Stamp Benefit Amounts According to SIPP and Administrative Records for Households With Nonzero Amounts in Both SIPP and the Records

Wave-Month	(N)	Mean Benefit Amount:		Mean Discrepancy	Std. Error of Mean	t	Percent Bias
		SIPP	Record				
1-4 mos. ago	(30)	\$102.00	112.40	-10.40	5.46	-1.90	-9.3%
1-3 " "	(32)	112.97	99.14	13.83	10.27	1.35	14.0%
1-2 " "	(29)	102.48	104.40	-1.92	4.00	-0.48	-1.8%
1-1 " "	(28)	99.39	102.03	-2.63	3.78	-0.70	-2.6%
2-4 " "	(24)	100.92	99.07	1.84	8.54	0.22	1.9%
2-3 " "	(21)	94.19	104.33	-10.14	6.03	-1.68	-9.7%
2-2 " "	(23)	98.48	97.96	0.52	4.32	0.12	0.5%
2-1 " "	(25)	89.84	84.57	5.27	5.58	0.94	6.2%

Table 7: Month-to-Month Food Stamp Participation Transitions According to SIPP and Administrative Records

Wave-Month Pair:	Transition Rate:		N	Mean Discrepancy	Std. Error of Mean	t	
	SIPP	Record					
1:	1-4 to 1-3	.0057	.0095	529	-.0038	.0053	-0.72
	1-3 to 1-2	.0057	.0076	529	-.0019	.0050	-0.38
	1-2 to 1-1	.0019	.0019	529	0	.0027	0
Seam:	1-1 to 2-4	.0195	.0058	513	.0136	.0070	1.94
2:	2-4 to 2-3	.0078	.0039	514	.0039	.0048	0.81
	2-3 to 2-2	.0096	.0191	523	-.0096	.0074	-1.30
	2-2 to 2-1	.0077	.0135	518	-.0058	.0064	-0.91

Table 8: Month-to-Month Food Stamps Benefit Amount Transitions According to SIPP and Administrative Records for Households With Nonzero Amounts in Both Months

Wave-Month Pair:	Transition Rate:		Mean Discrepancy	Std. Error of Mean	t	
	SIPP (N)	Record (N)				
1:	1-4 to 1-3	.2667 (30)	.6000 (35)	-.3333	.1233	-2.70
	1-3 to 1-2	.2414 (29)	.5152 (33)	-.2738	.1234	-2.22
	1-2 to 1-1	.2857 (28)	.3750 (32)	-.0893	.1217	-0.73
Seam:	1-1 to 2-4	.7500 (20)	.4138 (29)	.3362	.1434	2.34
2:	2-4 to 2-3	.0455 (22)	.4333 (30)	-.3879	.1198	-3.24
	2-3 to 2-2	.1905 (21)	.4167 (24)	-.2262	.1374	-1.65
	2-2 to 2-1	.3478 (23)	.3478 (23)	0	.1404	0