Exploring Social Security Payment History Matched with the Survey of Income and Program Participation

James Sears and Kalman Rupp
Social Security Administration
Office of Research, Evaluation, and Statistics[1]

Key words: survey error, administrative records, social security, SIPP

Since the pioneering 1973 CPS-IRS-SSA Exact Match Study (Kilss and Scheuren 1978), the use of linked data sets has been an important tool for analysts. Our current study is one of several recent efforts to improve understanding and usefulness of survey data matched to individual-level benefit records, and it is the first one involving actual payment histories from the Social Security Administration's (SSA's) Payment History Update System (PHUS) for Old Age, Survivors, and Disability Insurance (OASDI). The SSA administrative records have been matched to the Survey of Income and Program Participation (SIPP). In analyzing these matched data for elderly and working-age beneficiaries, we seek to identify specific factors primarily responsible for survey error and to assess the relevance of survey error for a range of different analytic objectives.

In this paper we address the following major issues:

- The availability of and limitations of SIPP matched to SSA records
- Errors arising from the past use of payment eligibility record data
- The magnitude of SIPP error in *receipt* of OASDI benefits measured against record data on actual receipt
- The magnitude of SIPP error in estimating *benefit amounts* conditional on receipt measured against record data on actual receipt
- Reasons for SIPP error and analytic implications

Finally, we comment on issues related to Social Security number (SSN) nonreporting, particularly in light of the fairly low 2001 SIPP SSN match rate. As part of our concluding comments we also reflect on the analytic issues that need to be jointly considered in the context of the twin issues of SIPP error versus SSN nonmatch error.

I. What matched SIPP and administrative data are available, and what are the limitations of their use?

All SIPP panels are matched to complete benefit histories for Supplemental Security Income (SSI) records and to OASDI benefit histories beginning for the year 1984[2]. However, matches are only made for individuals whose SSNs were provided during SIPP interviews. This limitation is due to privacy concerns rather than to any inability of the Census Bureau to determine SSNs by other means. Privacy concerns also dictate that the matched data may only be used by authorized researchers at SSA and Census Bureau locations. Although names and SSNs have been dropped, currently available matched files still contain too much sensitive individual information to allow their public use. Thus, a primary use of the matched records is to assess the validity of the unmatched SIPP data that remain the basis for many research projects.

We focus on the most recent SIPP panels because these will be of most interest to other researchers. Table 1 shows the SSN match rates at the start of the 1996 and 2001 panels. These are unweighted percentages, but the nearly identical weighted numbers are included in an appendix. The 1996 match rate of about 85 percent is very similar to match rates for earlier panels, while the 2001 match rate is only about 52 percent. Part of this difference is due to the fact that the 2001 SSNs were not collected until Wave 2, by which time about 15 percent of the panel had been dropped to reduce survey costs. However, even among the retained portion of the sample, the 2001 match rate is only about 60 percent. Researchers relying on matched data may wish to continue using earlier panels with higher match rates.

---

[1] The authors share equal responsibility for the paper. The address for both authors is SSA/ORES, 500 E Street, S.W., 9th Floor, Washington, D.C. 20254. Address correspondence to James Sears (202) 358-6104 (voice), (202) 358-6187 (fax). The authors' email addresses are jim.sears@ssa.gov and kalman.rupp@ssa.gov

[2] The SIPP data are also matched to earnings data from the Internal Revenue Service. See Vaughan and Scheuren (2002) for an example of research involving the matched earnings data.

Table 1: SSN Match Rates for the 1996 and 2001 SIPP Panels

	Ages 18 to 64	Ages 65 and older	All adults
March 1996	84.5%	85.7%	84.7%
January 2001, full sample	51.8%	50.6%	51.6%
January 2001, retained sample	60.5%	59.3%	60.3%

II. What is the magnitude of error arising from the past use of payment eligibility records, rather than records of actual payments?

While matched SSI payment records have been available for several years, virtually all previous analytic work using OASDI record data has been limited to "payment eligibility" rather than to actual payments. Should we now discard all these studies because now we have improved record data? It is not immediately obvious that a discrepancy between SSA payment eligibility record data and SIPP self-report is necessarily attributable to SIPP error. For example, disability applicants whose denial is reversed on appeal may be shown to be "eligible" to receive disability payments for several months prior to the reversal decision even though they have not actually received payments during this period; the actual payment is retroactive. This is not just a theoretical possibility. Therefore, it is no surprise that analysts in the past have been reluctant to label *discrepancies* between SIPP self-report and payment eligibility records as "SIPP error" (Olson 2001). In our previous study (Huynh, Rupp, and Sears 2001) we addressed this issue through indirect, aggregate evidence using data from SSA's Office of Quality Assurance (OQA), and concluded that at least half of the difference between the Monthly Benefit Payable (MBP) eligibility amount and SIPP self-report should be attributable to SIPP error, and perhaps substantially more. Therefore, with some caveats we accepted the SSA payment eligibility record data as an appropriate, albeit second best yardstick for measuring SIPP error.

Using the records on actual payments, we could test this conclusion against strong data. We find that the March 1996 mean average absolute error of the payment eligibility record data is about 2.5 percent when judged against the newly available actual payment record data, the error being slightly less than the lower-bound back-of-the envelope estimated range of 3 to 6 percent we inferred from OQA aggregate data in our previous paper (Huynh, Rupp, and Sears 2001). Our current estimate of the "true" mean average percent error attributable to the SIPP is about 10 percent, well within the 8.2-19.6 percent range posited in that paper. Our current estimates based on actual payment SSA record data indicate that the mean average SIPP error was 4 times larger than the mean MBP error, confirming our previous contention that the bulk of the estimated discrepancy between the imperfect MBP record data that were available to us at the time and the SIPP record is attributable to survey, rather than administrative record, error.

The corresponding January 2001 numbers indicate an even starker contrast. The mean absolute payment eligibility record error decreased to $12.30, while the corresponding SIPP mean absolute error increased to $98.80. This is not surprising, since more recent payment eligibility data should be a more accurate indicator of actual payments, simply because the shorter the period elapsing between the reference month and the record extraction date, the less likely it is that SSA will change the estimated payment eligibility. Importantly, the proportion of observations with large payment eligibility record (MBP) error (more than $200 in either direction) is small -- about 1 to 1.5 percent.

The two basic conclusions from this analysis are (a) previous analyses had the potential to substantially increase the precision of benefit data by replacing SIPP self reports with SSA payment eligibility record data (MBP) and (b) the use of actual payment record data further improves the potential of the record data to improve the accuracy of the matched data set. Because monthly MBP data are available for all years beginning with 1962 while PHUS data are first available for 1984, some continued use of the MBP may be warranted. Still, there are real improvements in accuracy from the use of actual payment record data.

III. What is the magnitude and pattern of SIPP error in receipt of Social Security benefits?

Each row of Table 2 contains the percentage breakdown of Social Security benefit types reported by SIPP respondents, providing comparable information to the Table 2 of Huynh, Rupp, and Sears (2001) but using the superior "actual payment" records from SSA for March 1996. In addition, we provide similar data for January 2001and include detail for elderly and working-age populations. The bold diagonal entries in Table 2 indicate percentages of matched SIPP respondents whose self

reports agree with administrative records. The table contains unweighted data and is limited to respondents who could be matched with administrative records. Comparable weighted numbers are provided as an Appendix A.

The overall findings confirm our previous results concerning confusion between OASDI and SSI as sources of payments. The January 2001 overall data are generally consistent with the March 1996 report, albeit the misreporting of the receipt of benefits from both SSA programs as "OASDI only" is clearly a much bigger problem. The differences between individuals 65 years old or older and working age beneficiaries (predominantly disabled worker beneficiaries) are notable. Reporting of the receipt of both types of benefits is more accurate among individuals 65 years old or older, particularly in 1996. The overall degree of misreporting receiving both benefits is larger for January 2001. This seems to be the result of a sharp increase in the percentage of individuals 65 years old or older incorrectly reporting OASDI only, a finding that cannot be explained away by the relatively small cell sizes. The reporting of OASDI only is clearly superior among individuals 65 years old or older as one might expect. The reporting of the nonreceipt of either benefit or the receipt of SSI only appears to be more accurate among working age beneficiaries. However, these findings should be interpreted in the context of the relative significance of SSI and the nonreceipt of either type of benefit for the two age groups.

Table 2: SIPP Reported Receipt of OASDI and SSI by SSA Payment Record Information

Month	Receipt based on SSA payment	SIPP report of receipt (%)				Total percent	N
		Both	Neither	SSI only	OASDI		
		Ages 18 to 64 only					
March 1996	Both	**72.58**	7.53	13.17	6.72	100.00	372
January 2001		**73.79**	6.31	11.65	8.25	100.00	206
March 1996	Neither	0.02	**99.49**	0.19	0.29	99.99	44,445
January 2001		0.04	**99.50**	0.16	0.31	100.01	26,197
March 1996	SSI only	3.04	9.11	**82.56**	5.28	99.99	757
January 2001		2.97	11.29	**78.42**	7.33	100.01	505
March 1996	OASDI only	1.39	9.21	3.08	**86.32**	100.00	2,302
January 2001		1.48	6.73	6.43	**85.37**	100.01	1,353
		Age 65 and older only					
March 1996	Both	**89.10**	0.44	1.78	8.67	99.99	450
January 2001		**76.92**	1.44	3.37	18.27	100.00	208
March 1996	Neither	0.48	**64.90**	0.80	33.81	99.99	624
January 2001		0.58	**70.23**	0.87	28.32	100.00	346
March 1996	SSI only	7.61	4.06	**71.07**	17.26	100.00	197
January 2001		1.87	4.67	**74.77**	18.69	100.00	107
March 1996	OASDI only	0.31	1.48	0.08	**98.13**	100.00	8,396
January 2001		0.30	1.72	0.11	**97.88**	100.01	4,706
		All adults					
March 1996	Both	**81.63**	3.65	6.93	7.79	100.00	822
January 2001		**75.36**	3.86	7.49	13.29	100.00	414
March 1996	Neither	0.03	**99.01**	0.20	0.76	100.00	45,069
January 2001		0.05	**99.11**	0.17	0.67	100.00	26,543
March 1996	SSI only	3.98	8.07	**80.19**	7.76	100.00	954
January 2001		2.78	10.13	**77.78**	9.31	100.00	612
March 1996	OASDI only	0.54	3.14	0.73	**95.59**	100.00	10,698
January 2001		0.56	2.84	1.52	**95.08**	100.00	6,059

IV. What is the magnitude of SIPP error in estimating benefit amounts conditional on receipt?

Table 3 provides percent distributions of SIPP errors, using the more accurate "actual payment" records. This unweighted tabulation is limited to people who were beneficiaries according to both administrative and SIPP data[3]. Thus it is intended to assess the validity of self-reported benefit amounts among people who accurately indicate their receipt of benefits. As in Table 2, data from the first waves of the 1996 and 2001 SIPP panels are presented separately for aged and disabled respondents.

The March 1996 results are quite close to the estimated errors previously derived using the less accurate MBP administrative eligibility yardstick (Huynh, Rupp, and Sears 2001). The percentage of March 1996 respondents who report the administrative OASDI benefit exactly improves to 51 percent with the PHUS data from 46 percent in the earlier MBP study, but there is no corresponding improvement in estimated mean SIPP errors. Also, the proportion of high and low outliers is approximately as high as previously seen with the MBP. For example, 10.9 percent of adults underreport the March 1996 PHUS benefit amount by more than $100, and 10.9 percent also underreport the March 1996 MBP amount by more than $100. This suggests that most previously observed large outliers are attributable to SIPP, rather than payment eligibility record error.

More troubling is that reporting of OASDI amounts appears generally less accurate in 2001 than in 1996.[4] The mean absolute errors and percentages of outliers increase for both age groups, but the changes are most dramatic for the elderly. Nonetheless, the overall shape of the distributions is rather close for the two groups of beneficiaries in both years.

Table 3: Magnitude of SIPP Error as Measured by SSA Payment Record Information

SIPP minus SSA record amount (rounded to nearest $)	OASDI net of SMI premium (percent distributions)					
	Ages 18 to 64		Ages 65 and older		All adults	
	March 1996	January 2001	March 1996	January 2001	March 1996	January 2001
Less than -200	7.1	7.9	7.3	11.2	7.3	10.5
-200 to -101	2.9	3.9	3.8	4.1	3.6	4.0
-100 to -51	3.3	4.4	4.8	5.5	4.5	5.2
-50 to -41	2.6	1.9	1.7	2.1	1.9	2.0
-40 to -31	1.4	2.6	1.6	2.9	1.5	2.8
-30 to -21	3.3	3.4	3.6	3.8	3.5	3.7
-20 to -11	3.2	5.1	3.5	4	3.4	4.3
-10 to -1	8.7	7.1	6.3	6.3	6.8	6.5
zero	52.6	44.1	50.7	40.8	51.1	41.5
1 to 10	4.0	5.5	4.0	3.8	4.0	4.2
11 to 20	1.4	1.1	1.5	1.6	1.5	1.5
21 to 30	0.8	1.5	1.3	1.1	1.2	1.2
31 to 40	0.9	1.0	1.1	1.2	1.1	1.1
41 to 50	1.1	1.2	1.3	1.9	1.3	1.7
51 to 100	2.4	2.3	2.1	2.9	2.2	2.8
101 to 200	1.6	2.7	1.9	2.1	1.8	2.2
over 200	2.8	4.2	3.5	4.9	3.3	4.7
N	2,314	1,344	8,705	4818	11,019	6,162
Est mean SIPP error ($)	-41.2	-38.5	-24.1	-35.7	-27.7	-36.3
Mean absolute SIPP error ($)	81.8	88.5	70.5	101.6	72.8	98.8

[3] See the appendix for a weighted version of this tabulation.

[4] Reports of SSI benefit amounts would show a similar drop in accuracy between 1996 and 2001. For example, among adults who correctly reported receiving SSI, 57.0 percent exactly reported their correct benefit amounts for March 1996, while only 41.5 percent reported correct benefit amounts for January 2001.

V. What are the reasons? What are the analytic implications?

It is useful to think about possible reasons for SIPP error along two dimensions – cross-section and longitudinal – with survey respondents and the survey itself as the main culprits (obviously, interaction effects may be relevant as well). Cross-sectionally, one might think of different groups of people in terms of factors such as receipt of OASDI and/or SSI benefits, participation in Part B of Medicare, stability of life patterns, cognitive abilities, mental health, presence of a "representative payee", and availability and willingness to talk to interviewers on the phone or in person. Survey characteristics may include factors such as mode (phone, in-person or mixed), interviewer skill and aptitudes, survey policies concerning eligible respondents, resources devoted to tracking and refusal conversion, incentive payments, and statistical procedures to deal with incomplete, missing or dubious self-reports. These two sets of factors may interact in producing differences in survey error by interview and imputation status, important cross-sectional factors we identified in our previous paper.

The longitudinal dimension is just as important because both the SIPP and the SSA record data are longitudinal (See Vaughan and Scheuren, 2002, for pioneering work in this area). Survey attrition is affected both by population characteristics such as migration or institutionalization and by survey methods such as tracking, locating, policies concerning people who move outside of the survey PSUs, and forms of longitudinal imputation. We believe that patterns of benefit receipt over time may also affect respondent error. Not only may concurrent receipt of OASDI and SSI be a source of respondent error, but transitions within the two programs may cause confusion. For example, the transition from SSI to OASDI at the end of the DI waiting period might be misreported. Other sources of benefit changes that might cause survey errors for particular individuals include the death of a spouse, large lump-sum payments following the long adjudication of disability cases, and agency errors resulting in erratic payment patterns. Likewise, the duration of benefit receipt should be relevant here, particularly in proxy respondent situations. Survey error may take the form of temporal misalignment of the receipt of benefits with the correct total amount of money being reported for an incorrect period. Any benefit change could increase survey errors, even if it is across-the-board in nature, such as annual cost of living adjustment. COLAs may be a source of survey error for several reasons, such as underreporting due to lags in respondent perceptions (possibly compounded by electronic benefit payments) and overreporting for the months preceding a COLA due to the well-known SIPP seam effect. The administrative record data – especially data on actual payments – are particularly useful because they are largely unaffected by these various longitudinal sources of error, and they should also be useful in modeling SIPP error.

Table 4 addresses the COLA and seam effect issues by comparing Wave 1 SIPP errors for December 2000 with Wave 2 errors for May 2001. COLAs are received each year beginning in January. The December reports should not include the 2001 COLA, but they were collected after beneficiaries had begun receiving the COLA. The changing benefit amount is most likely to blame for the fact that under 24 percent of December self-reported amounts are exactly correct, versus more than 40 percent of the May amounts. Although this is not shown in the table, less than half of the Wave 1 OASDI beneficiaries reported any change in benefit amount between December and January. Many appear to have reported the COLA for all months in Wave 1, but a few first report increased benefits for the Wave 2 months. The COLA issue may also go part way toward explaining the apparent decrease in accuracy over time that was observed in Table 3. While only about 40 percent of beneficiaries report exactly correct benefits for May 2001, an additional 7 percent over-report by only one dollar. If one takes into account this possible rounding error, the May 2001 self reports appear more accurate but still not quite as accurate as the March 1996 self reports shown in Table 3.

Table 4: SIPP Errors before and after COLA

SIPP minus SSA record amount (rounded to nearest $)	OASDI net of SMI	
	All adults	
	Dec. 2000	May 2001
Less than -200	10.6	9.0
-200 to -101	4.5	3.9
-100 to -51	5.1	5.6
-50 to -41	1.7	1.8
-40 to -31	2.0	2.4
-30 to -21	2.7	3.0
-20 to -11	3.9	3.8
-10 to -1	7.0	6.3
zero	23.9	40.4
1 to 10	8.3	10.5
11 to 20	7.1	1.8
21 to 30	6.0	1.1
31 to 40	3.9	1.1
41 to 50	2.3	1.5
51 to 100	3.5	2.4
101 to 200	2.7	1.7
over 200	4.8	3.7
N	4,543	6,233
Est mean SIPP error ($)	-28.9	-32.5
Mean absolute SIPP error ($)	108.9	86.2

The SSN match issue is also relevant to the apparent decline in accuracy. The lack of a valid SSN match can undermine the potential benefits of using matched data in two fundamental ways. First, it reduces the effective sample size available for the analyst, unless a decision is made to use a combination of matched and unmatched records. Second, if the SSN match is systematically related to observable, and especially

unobservable characteristics of individuals on factors relevant to the purpose of any given piece of analysis, it may introduce systematic bias in analyses restricted to the matched data. It is important to realize, however, that this is a problem only to the extent that the nonmatch is nonrandom with respect to the phenomenon to be studied. It may even be the case that a reduction in the match rate would reduce the mean of systematic differences between the matched and nonmatched cases. This might occur if the most likely non-match cases are very different from the rest (e.g., immigrants), while additional attrition is less selective. For example, the additional attrition could be related to survey shortcomings or changes in population tastes for sharing sensitive SSN information that are largely uncorrelated with the various relevant characteristics of the people in the sample. On the other hand, if the proportion of immigrants changed substantially over time, the associated attrition bias might increase.

In assessing the analytic implications, both survey error and the quality of the SSN nonmatch need to be considered carefully. Most importantly, perhaps, the analytic objectives and techniques used to compensate for either or both are key to answering the questions: Does it matter? How much? A number of analytic studies have shown that if sophisticated analytic techniques are used the implications of the errors for the quality of conclusions may not be as important as the naïve observer may think (See Zabel, 1998 for an illuminating analysis of PSID and SIPP attrition). However, a data set that may be acceptable for some purposes may be regarded as fatally flawed for others. In particular, if the analytic objective focuses on small subgroups particularly prone to survey or SSN match error the data set may turn out to be very unreliable, while it may be very reliable for the study of large population groups (the examples of studying small subgroups of disabled SSI beneficiaries versus all working aged adults come to mind). Finally, a data set may be appropriate for descriptive purposes, while inappropriate for causal modeling or vice versa.

An important issue when using the matched data set is the unit of observation to be used in establishing a "match" or "nonmatch" situation. In previous discussions of the SSN match issue the focus typically has been on SSN matches at the individual level – the presence or absence of a match defined with respect to a given reference person. However, in some studies families and household relationships are of inherent interest. For example, in the context of multirecipient SSI households and families Koenig and Rupp (2002) used a more restrictive match screen: all household or family members (depending on the specific focus of a piece of analysis) had to have a "valid" SSN match. This kind of decision obviously has costs: the rule is bound to reduce the match rate; thus tradeoffs need to be considered by the analyst.

Another important issue with respect to the use of matched data sets is whether to base the analysis entirely on the survey subuniverse with valid matches, or to combine survey and record information for matched and non-matched cases, perhaps with some aggregate adjustment accounting for non-matched cases. In some cases the decision is straightforward: if one wants to study long term earnings or benefit histories, there is no practical way of imputing meaningful values for the non-matched cases. However, in other cases the benefits of combining the matched and non-matched subuniverses outweigh the disadvantages. This may be the typical situation when the subsample of interest is very small to start with.

The one important thing to keep in mind is that the matching of survey and administrative records creates a practically no-lose situation Once a matched data set is created, it is virtually costless to use it for any analytic purpose, whether the analytic gains are small or large in the given case.

References

Huynh, Minh, Kalman Rupp, and James Sears (2002), "The Assessment of Survey of Income and Program Participation Benefit Data Using Longitudinal Administrative Records," The Survey of Income and Program Participation Report No. 238. Washington, D.C.: U.S. Department of Commerce, Census Bureau.

Kilss, Beth and Fredrick J. Scheuren (1978), "The 1973 CPS-IRS-SSA Exact Match Study," *Social Security Bulletin* 41(10): 14-22.

Koenig, Melissa and Kalman Rupp (2002), "The Poverty Status of Different Types of Multirecipient Households: Is SSI Fair to Married Couples?" presented at the Annual Meeting of the Population Association of America, with a revised version forthcoming in the *Social Security Bulletin*.

Olson, Janice A. (2001), "Social Security Benefit Reporting in the Survey of Income and Program Participation and in Social Security Administrative Records," The Survey of Income and Program Participation Report No. 235. Washington, D.C.: U.S. Department of Commerce, Census Bureau.

Vaughan, Denton and Fritz Scheuren (2002), "Longitudinal Attrition in SIPP and SPD," presented at the American Statistcial Association's Joint Statistical Meetings.

Zabel, Jeffery E. (1998), "An Analysis of Attrition in the Panel Study of Income Dynamics and the Survey of Income and Program Participation with an Application to a Model of Labor Market Behavior," *The Journal of Human Resources* 33(2): 479-506.

Appendix A

This set of weighted tables provides comparable numbers to the unweighted text tables. While study of SIPP errors begins with the examination of individual, unweighted cases, we want to be sure that our conclusions also apply for weighted SIPP populations. The weighted tabulations here are qualitatively similar to the unweighted numbers in all cases.

Table A1: SSN Match Rates for the 1996 and 2001 SIPP Panels, Weighted

	Ages 18 to 64	Ages 65 and older	All adults
March 1996	84.5%	85.7%	84.7%
January 2001, full sample	51.8%	50.7%	51.6%
January 2001, retained sample	60.4%	59.3%	60.2%

Table A2: SIPP Reported Receipt of OASDI and SSI by SSA Payment Record Information, Weighted

Month	Receipt based on SSA payment	SIPP report of receipt (%)				Total percent	unweighted N
		Both	Neither	SSI only	OASDI		
				Ages 18 to 64 only			
March 1996	Both	**71.22**	7.41	14.73	6.64	100.00	372
January 2001		**71.97**	7.37	12.41	8.25	100.00	206
March 1996	Neither	0.03	**99.49**	0.20	0.28	100.00	44,445
January 2001		0.04	**99.52**	0.13	0.30	99.99	26,197
March 1996	SSI only	3.02	9.45	**82.14**	5.38	99.99	757
January 2001		2.40	11.52	**78.75**	7.33	100.00	505
March 1996	OASDI only	1.43	9.03	3.22	**86.32**	100.00	2,302
January 2001		1.35	6.88	6.26	**85.51**	100.00	1,353
				Age 65 and older only			
March 1996	Both	**88.30**	0.41	1.71	9.58	100.00	450
January 2001		**77.66**	1.38	2.65	18.31	100.00	208
March 1996	Neither	0.37	**65.76**	0.66	33.21	100.00	624
January 2001		0.51	**69.97**	0.64	28.89	100.01	346
March 1996	SSI only	7.86	4.54	**70.14**	17.46	100.00	197
January 2001		2.26	2.06	**76.92**	18.77	100.01	107
March 1996	OASDI only	0.31	1.46	0.09	**98.14**	100.00	8,396
January 2001		0.29	1.59	0.08	**98.04**	100.00	4,706
				All adults			
March 1996	Both	**79.96**	3.83	8.07	8.14	100.00	822
January 2001		**74.72**	4.47	7.68	13.12	99.99	414
March 1996	Neither	0.03	**99.03**	0.20	0.73	99.99	45,069
January 2001		0.05	**99.14**	0.14	0.68	100.01	26,543
March 1996	SSI only	4.05	8.41	**79.60**	7.94	100.00	954
January 2001		2.37	9.84	**78.42**	9.36	99.99	612
March 1996	OASDI only	0.54	3.04	0.74	**95.68**	100.00	10,698
January 2001		0.52	2.74	1.42	**95.31**	99.99	6,059

Table A3: Magnitude of SIPP Error as Measured by SSA Payment Record Information, Weighted

SIPP minus SSA record amount (rounded to nearest $)	OASDI net of SMI premium (percent distributions)					
	Ages 18 to 64		Ages 65 and older		All adults	
	March 1996	January 2001	March 1996	January 2001	March 1996	January 2001
Less than -200	6.9	7.8	7.4	11.5	7.3	10.7
-200 to -101	2.8	4.2	3.8	4.1	3.6	4.1
-100 to -51	3.3	4.3	4.9	5.6	4.6	5.3
-50 to -41	2.7	1.9	1.7	2.1	1.9	2.1
-40 to -31	1.4	2.8	1.5	2.9	1.5	2.9
-30 to -21	3.5	3.6	3.5	3.8	3.5	3.8
-20 to -11	3.0	4.8	3.5	3.9	3.4	4.1
-10 to -1	8.7	7.2	6.3	6	6.7	6.3
zero	52.1	43.6	50.5	40.6	50.8	41.3
1 to 10	3.9	5.5	4.0	3.8	4.0	4.2
11 to 20	2.0	1.2	1.5	1.6	1.6	1.5
21 to 30	0.9	1.6	1.3	1.0	1.2	1.1
31 to 40	0.9	1.0	1.1	1.1	1.1	1.1
41 to 50	1.1	1.2	1.4	1.8	1.3	1.7
51 to 100	2.5	2.3	2.2	3.1	2.3	2.9
101 to 200	1.6	2.6	1.9	2.1	1.8	2.2
over 200	2.7	4.3	3.5	4.8	3.4	4.7
Unweighted N	2,314	1,344	8,705	4818	11,019	6,162
Est mean SIPP error ($)	-42.7	-40.4	-23.8	-37.3	-27.7	-37.9
Mean absolute SIPP error ($)	83.7	92.6	71.0	103.6	73.6	101.3

Table 4A: SIPP Errors before and after COLA

SIPP minus SSA record amount (rounded to nearest $)	OASDI net of SMI	
	All adults	
	Dec. 2000	May 2001
Less than -200	10.8	9.3
-200 to -101	4.5	3.9
-100 to -51	5.1	5.6
-50 to -41	1.7	1.8
-40 to -31	2.1	2.4
-30 to -21	2.7	3.0
-20 to -11	4.1	3.7
-10 to -1	6.8	6.1
zero	23.7	40.4
1 to 10	8.1	10.5
11 to 20	7.0	1.8
21 to 30	6.0	1.1
31 to 40	3.8	1.0
41 to 50	2.4	1.6
51 to 100	3.7	2.4
101 to 200	2.9	1.7
over 200	4.7	3.7
Unweighted N	4,543	6,233
Est mean SIPP error ($)	-28.3	-34.2
Mean absolute SIPP error ($)	113.1	87.0