The assessment of Survey of Income and Program Participation (SIPP) benefit data using longitudinal administrative records*

Minh Huynh, Kalman Rupp, and James Sears
Office of Research, Evaluation, and Statistics, Social Security Administration

Abstract

Using Social Security Administration administrative records matched to the 1993 and 1996 SIPP panels, we assess discrepancies in reports of benefit receipt and benefit amounts for four sample months. The Supplemental Security Income (SSI) "payment" records are excellent sources of "true" Federally administered SSI payments. The Old-Age, Survivors, and Disability Insurance (OASDI) "payment eligibility" records used in this analysis are imperfect proxies for "true" OASDI payments, but we provide evidence suggesting that over half of the observed absolute error is attributable to the SIPP, and the true SIPP error may even be larger than the discrepancy we observe. We present evidence of confusion between the OASDI and SSI as sources of payments, and we observe a higher incidence of reporting errors for SSI than for OASDI. After wave 1 of the 1993 SIPP, survey responses tend to reflect OASDI benefits net of the Medicare Supplementary Medical Insurance (SMI) premium. While this is consistent with revised questionnaire wording, the use of survey data without adjusting for the SMI premium could substantially bias SIPP-based estimates of total income and poverty status. We find that reporting errors for both SSI and OASDI differ dramatically by imputation status and provide evidence that errors may be systematically related to sample attrition and interview status (self, proxy, and refusal). Finally, we provide a brief assessment of the effect of the lack of social security numbers in a nontrivial fraction of cases and find clear evidence of selectivity.

Keywords: survey error, administrative records, Social Security, Supplemental Security Income, Survey of Income and Program Participation

This paper uses administrative records data from the Social Security Administration (SSA) to assess the accuracy of SIPP data concerning Old-Age, Survivors and Disability Insurance (OASDI) and Supplemental Security Income (SSI) benefits. Our findings concern (a) aggregate benefit amounts; (b) receipt of benefits; (c) per recipient benefit amounts; (d) attrition, nonresponse, proxy and imputation error; (e) the treatment of the Supplementary Medical Insurance (SMI) premium in SIPP data; and (f) selectivity in reporting SSNs.

Without attempting to provide a comprehensive review of the literature, we note that recent work has focused on the accuracy of administrative records (Pickett and Scott 1996, Rupp et al 1999), on comparisons of SIPP and administrative record earnings information (Pedace and Bates 2001), on comparisons of Master Beneficiary Record (MBR) data to 1990 SIPP information (Olson 2001) and on analytic strategies for combining survey and administrative records data (Rupp and Davies 2000). The principal contributions of this paper include the joint examination of Social Security and SSI information, comparisons involving multiple months and panels

* The authors share equal responsibility for this paper and would like to thank the following people for providing information and feedback: Mary Barbour, Pat Doyle, Howard Iams, Melissa Koenig, Gordon Lester, Jan Olson, Ken Olson, Charlie Scott, and Denny Vaughan. The views expressed here are those of the authors and do not represent the official position of the Social Security Administration or any other federal agency. Please address correspondence to: Kalman.Rupp@ssa.gov.

(1993 and 1996 panels), and the use of longitudinal administrative records to examine attrition, nonresponse and imputations.

I Data

Survey and administrative record data may differ for three main reasons: (1) administrative record error; (2) survey error; and (3) matching error. While benefit data from SSA's well-established payment-oriented records system are widely regarded to be more reliable than survey data, any comparison should consider all possible sources of error. This is especially true with respect to the SIPP, which collects detailed monthly information on benefit receipt and amounts and which is believed to provide high quality data in this area.

The MBR and Supplemental Security Record (SSR) data systems contain complete monthly benefit information for all persons who ever received OASDI or SSI benefits. These data are unaffected by attrition, refusal, or misreporting by self or proxies. They cover the complete history of benefits received from birth to record extraction or death, whichever occurs first. These features make the MBR and SSR attractive tools for analyzing SIPP attrition and nonresponse and for assessing the accuracy of SIPP reported benefits.

The MBR and SSR do have important limitations. The most notable limitation of currently available MBR data – Monthly Benefit Payable (MBP) and Monthly Benefit Credited (MBC) variables – is that they reflect payment eligibility rather than actual payments. This means that the MBR provides the amount SSA believes it should have paid rather than the amount it actually did pay during a given month. Based on data from SSA's Office of Quality Assurance (OQA), we estimate that the mean absolute error in the MBP as a measure of payments sent is between 3 percent and 6 percent[1]. While the SSR does provide information on the actual benefit check or electronic payment amounts sent out by the Agency each month, it excludes State-administered supplementation. State-administered supplementation comprised about 21 percent of all state SSI supplements in 1999. Neither the MBR nor the SSR accounts for benefit payments lost due to factors such as incorrect address information or lost checks. While acknowledging that the administrative record data are not "perfect" due to these and other factors, we regard the SSA record data to be useful in assessing SIPP error. We regard the SSR "payment" amount information to be extremely accurate because the source is the system that is used to execute actual payment actions. Our analysis using the OQA data indicates that the SIPP error in measuring OASDI, which we cannot directly observe, is over half of the "error" we do observe, and it may even be larger than the discrepancy we observe. For January 1993 we see a mean absolute difference of 13.9 percent between SIPP and MBP, we estimate that the "true" mean absolute SIPP error is bounded by 8.2 and 19.6 percent.

[1] This estimate relies on work by Ken Olson of OQA in response to a request by Jan Olson (2001). Ken Olson used a sample of 624,789 payment transactions for the year 2000 to estimate the dollar amount of corrections made to account for previous "underpayments" and "overpayments." We agree with Jan Olson (2001) that this information, while suggestive, is not conclusive about the incidence of per beneficiary receipt measurement error. However, the OQA data do provide unambiguous results concerning the aggregate amount of both "underpayments" (about $785 million per month) and "overpayments" (about $159 million per month). We used these results to calculate bounds on the mean absolute error in the "Monthly Benefit Paid" variable as a measure of OASDI payments (net of the SMI premium). Detail is available from the authors.

The most recent available SIPP data come from the 1996 panel, but it does not yet contain longitudinal edits that will be made for the full-panel file. Also, matched administrative records are not yet available for individuals who entered the 1996 SIPP after the first wave.[2] We, therefore, rely on both the 1993 and 1996 panels. This approach allows us to assess whether findings are consistent across SIPP panels. While matched records are longitudinal, for presentational purposes we focus on four sample months. These include the first month for which interview data are available for all 4 rotation groups (January 1993 and March 1996) and the last month with similarly complete data for the 1993 panel (August 1995). Although the 1996 panel covers a longer span, we chose October 1998 so that the observation points would span the same amount of time (32 months) for both panels. Our SIPP analyses rely on unweighted data for individuals age 18 and over.

While the MBR and SSR are our major sources of administrative record data, the Summary Earnings Record (SER) provides the universe of social security numbers (SSNs) for matching with the SIPP. Not all SIPP respondents provided valid SSNs. Among our analysis group, approximately 87 percent of person records (32,510 individuals) in the SIPP have valid SSN matches at the start of the 1993 panel. The comparable figure is 84 percent (42,771 individuals) for the 1996 panel. In some cases more than one SIPP respondent was matched to the same SSN. We exclude all such individuals (516 in 1993 and 472 in 1996) from our count of matched records.

Even if both administrative records and the survey information were completely accurate matching error might still cause discrepancies. This is obviously the case when records are incorrectly merged due to incorrect identifiers (e.g. SSNs), but discrepancies may also arise from conceptual differences. For example, one source may reflect the month for which benefits are paid, while the other the month in which payments are received. Likewise, the units receiving the payment may be differently defined, such as individual versus couple or differences in the definition of family and household. We considered many potential sources of matching error and made corrections and edits (e.g. missing values versus zeros, considering information on "current pay status" in conjunction with benefit amounts) as appropriate. These tests and edits revealed no evidence of major additional sources of matching error. Importantly, we tested the accuracy of the SSN match by conducting a sensitivity analysis considering birth information, and found that the results presented below are robust to the SSN match.

II Total benefits

A basic comparison between the two data sets involves the relationship between the total amount of benefits derived from the SIPP and the SSA records. In this paper we use unweighted data. As shown in Table 1, for January 1993 the SIPP overestimates the aggregate amounts of OASDI (compared to MBP) and SSI benefits by 6.2 percent and 4.0 percent respectively. Much of the OASDI error can be attributed to reports of the SMI premium during the first wave of the 1993 SIPP. If, instead of comparing SIPP OASDI reports to the MBP, we use the gross MBC, the aggregate OASDI error drops from 6.2 percent to 1.4 percent. The estimated SSI error increases

[2] This limitation makes the SSN match selectivity problem worse for observations beyond the first wave of the 1996 panel. Matching for people who move into the sample during subsequent waves is currently underway at SSA. Updated matched data will be available in the future.

substantially for August 1995, while the OASDI error (net of the SMI premium) drops. The August 1995 figures are 22.9 percent and –1.1 percent for SSI and OASDI respectively. The 1996 panel data represent a more marked SIPP underestimate of total MBP amounts (3.2 percent for March 1996 and 3.7 percent for October 1998). Total SSI benefits are underestimated by the SIPP for March 1996 (4.5 percent) and for October 1998 (8.4 percent). All of these numbers are statistically different from zero at the 95 percent level. The OASDI estimates from the SIPP are consistently and substantially lower than the MBC estimates of gross OASDI benefits (6-8 percent difference). We note that the omission of SMI premium amounts from SIPP reports of OASDI benefits – if unadjusted – leads to nontrivial downward bias in estimating total income and upward bias in SIPP-based estimates of poverty. In 1998 the usual SMI premium was $43.80, a substantial portion of monthly income for the poor. According to the 2000 report of the Board of Trustees of the Federal Supplementary Medical Insurance Trust Fund, the aggregate amount of SMI premium payments by enrollees amounted to about $20 billion in 1998.

Table 1: Aggregate differences in benefits between SIPP and administrative records

Percent difference SIPP - SSA	SSI SIPP - Total from SSR	OASDI (net of SMI premium) SIPP - MBP	OASDI (gross) SIPP - MBC
January 1993	4.0	6.2	1.4
August 1995	22.9	-1.1	-6.3
March 1996	-4.5	-3.2	-7.9
October 1998	-8.4	-3.7	-8.2

The aggregate amount of benefits is the product of the number of persons in the universe, the probability of receipt and the mean amount of benefits conditional on receipt. Thus error in *aggregate benefit amount* may be due to mismeasuring *receipt* or mismeasuring *benefits conditional on receipt*. The following sections focus on (a) receipt and (b) benefit amounts conditional on receipt.

III Receipt of benefits

We start our analysis of receipt with aggregate SSA-SIPP comparisons, using the MBP variable from the MBR and combined Federal and State amounts from the SSR. Overall, the number receiving OASDI is overestimated by 2 percent in January 1993, and by 1.2 percent in August 1995. SSI receipt is overestimated in January 1993 by 4.2 percent and by 12.8 percent in August 1995. The SIPP and administrative records OASDI receipt probability averages are essentially identical for the two 1996 panel observations (-0.6 percent and 0.1 percent difference for the two observation points). However, both the March 1996 and October 1998 SIPP underestimate aggregate SSI receipt (by 4.5 percent and 1.8 percent, respectively). The contrast between OASDI and SSI overall receipt accuracy is striking. All of the 4 OASDI aggregate errors are within 2 percent. Substantial SSI overestimates in the 1993 panel give way to underestimates in the 1996 panel.

Next we look at the individual-level variation beyond these overall measures of SIPP receipt error. Table 2 shows SIPP-reported receipt for the possible combinations of OASDI and SSI receipt based on SSA records. Overall, the accuracy of reporting receipt of “OASDI only” or “neither” is very high. We note, however, that the seemingly small fraction of those whose “true”

receipt is OASDI or "neither" and incorrectly report SSI is relatively high in the SSI context. This is so because the prevalence of SSI receipt is much smaller than the prevalence of "OASDI only" and receiving "neither" benefit. The percent misreporting in the two categories involving SSI receipt is much higher. A nontrivial fraction of those receiving SSI ("SSI only" and "concurrent" SSI and OASDI) according to SSA records are misclassified in the SIPP as receiving "OASDI only". This portion is especially high among concurrent recipients. Misreporting in the opposite direction is much less common, but still nontrivial for the "concurrent" category. Finally, a substantial portion of "SSI only" recipients report no benefit at all. These patterns are not surprising given the presumed stigma associated with SSI, the greater visibility of OASDI, the fact that the vast majority of SSI recipients receive benefits administered exclusively by SSA, and that the receipt of SSI often precedes the receipt of OASDI among disability applicants for a few months. The patterns are fairly consistent for the four time points. Interestingly, for both panels, accuracy goes up between the two observation points for the two categories involving SSI receipt.

Table 2: SIPP report of OASDI and SSI benefit receipt by administrative record of receipt for adults with matched SIPP records

Year	Month	MBR and SSR receipt	SIPP report of receipt (%)				Total percent	N
			Both	Neither	SSI only	OASDI only		
1993	January	Both	**76.08**	3.49	5.91	14.52	100.00	372
1995	August		**80.75**	2.48	5.90	10.87	100.00	322
1996	March		**74.71**	4.89	7.99	12.40	99.99	613
1998	October		**80.06**	3.81	4.11	12.02	100.00	341
1993	January	Neither	0.06	**98.32**	0.37	1.25	100.00	25,704
1995	August		0.07	**97.99**	0.50	1.44	100.00	22,436
1996	March		0.04	**98.81**	0.17	0.97	99.99	33,545
1998	October		0.05	**98.66**	0.23	1.07	100.01	16,677
1993	January	SSI only	6.01	6.56	**78.69**	8.74	100.00	366
1995	August		3.60	9.14	**81.16**	6.09	99.99	361
1996	March		4.81	8.94	**78.54**	7.70	99.99	727
1998	October		3.02	9.32	**79.85**	7.81	100.00	397
1993	January	OASDI only	0.30	3.38	0.38	**95.95**	100.00	6,068
1995	August		0.37	4.35	0.55	**94.73**	100.00	5,632
1996	March		0.41	4.31	0.82	**94.46**	100.00	7,886
1998	October		0.65	3.77	0.81	**94.78**	100.01	4,328

IV Benefit amounts conditional on receipt

The MBR contains three different measures of monthly OASDI benefits. For this analysis we decided to use the MBP, which nets out the SMI premium for those who pay this Medicare Part B premium – the majority of OASDI beneficiaries. We use the MBP for consistency with the SIPP questionnaire, even though the SMI premium amount is conceptually and legally part of the person's OASDI income. Because SSA automatically deducts the SMI premium from the amount of benefits the person is entitled to, it is this net amount (OASDI benefit less any SMI premium) that is mailed or electronically transferred to beneficiaries.

The primary SSI income amount on the SIPP refers to "Federal supplemental security income (SSI)[3]." We compare this amount with the combined Federal and (Federally administered) State supplementation amount from the SSR. When we experimented with using only the Federal SSI payment amount variable from the SSR as the basis of comparison, the SIPP appeared to overestimate SSI benefits by an amount roughly equal to Federally administered *State* benefits. Therefore, we inferred that the main SIPP SSI benefit variable in practical terms best represents *Federally administered SSI benefits* that include *both* Federal benefits and State supplements. Our interpretation is supported by the fact that both the Federal payment and the State supplementation are issued by SSA as a *single* check or electronic transfer.

Table 3 presents the unweighted distribution of error in the monthly amount of OASDI and SSI benefit recorded. The table is limited to persons with a *positive receipt indicator* in both data sets. The results are striking. In January 1993 a large plurality (42.5 percent of observations) have OASDI benefit amounts that exceed the MBP by $31 to $40. For each of the other 3

Table 3: Distribution of discrepancies in benefit amounts for adults with positive amounts in both SIPP and administrative records

SIPP minus SSA record amount (rounded to nearest $)	OASDI (net of SMI premium)				Supplemental Security Income			
	Jan 93	Aug 95	March 96	Oct 98	Jan 93	Aug 95	March 96	Oct 98
	Percent distribution							
less than - 200	3.0	4.9	7.2	6.6	2.6	2.6	3.1	3.1
-200 to -101	2.1	3.2	3.7	4.3	1.1	1.2	1.7	2.6
-100 to -51	2.8	4.9	4.9	5.8	2.3	2.7	2.6	2.3
-50 to -41	1.0	1.8	2.4	2.3	0.7	1.4	0.7	1.5
-40 to -31	0.8	2.5	1.9	2.7	1.0	1.0	0.6	1.0
-30 to -21	1.5	3.8	3.8	4.0	1.3	1.2	2.0	2.6
-20 to -11	2.2	5.8	4.2	5.6	4.9	4.6	3.9	2.9
-10 to -1	3.1	11.2	9.0	10.7	19.8	14.9	13.2	17.4
zero	**7.6**	**39.2**	**46.3**	**35.3**	**51.7**	**55.7**	**56.5**	**39.8**
1 to 10	2.8	5.4	3.8	9.5	7.6	5.3	5.7	14.0
11 to 20	4.1	2.3	1.5	2.7	1.0	2.1	0.9	2.1
21 to 30	5.9	2.0	1.2	1.3	0.3	0.7	0.5	0.8
31 to 40	**42.5**	1.8	1.1	1.2	1.0	0.7	0.2	1.5
41 to 50	2.5	3.9	1.2	1.4	0.5	0.5	0.5	0.3
51 to 100	12.3	2.5	2.2	2.1	1.1	1.4	2.1	3.6
101 to 200	2.5	1.9	1.9	2.0	0.5	1.2	1.3	2.1
over 200	3.1	2.8	3.7	2.6	2.6	2.9	4.4	2.6
Total percent	99.8	99.9	100.0	100.1	100.0	100.1	99.9	100.2
N	6,177	5,651	8,015	4,444	615	585	1,113	616
Mean error	$26.6	-$13.0	-$17.3	-$26.4	-$17.7	-$1.1	-$0.4	-$16.2
Mean absolute error	$64.5	$52.8	$68.7	$61.0	$42.2	$27.4	$43.2	$47.9

[3] While individuals may be eligible for State SSI supplementation even if they do not receive Federal benefits, the SIPP only asks a second question about State supplementation amounts of people who have already reported Federally-administered benefits. We considered this question out of concern that respondents might not distinguish between State administered and Federally administered forms of State supplementation. While, the 1993 SIPP aggregated the amount from this State supplementation question with "black lung/state temporary disability benefits/Indian, Cuban or refugee assistance/national guard or reserve forces retirement," the 1996 panel does provide a separate State supplementation variable. This variable has positive values for 0.2 percent of our analytic sample or about 6 percent of the SSR-reported SSI beneficiary observations for March 1996. We note that this would not be sufficient to account for Federally administered State supplementation.

observations, less than 2 percent of individuals fall in the equivalent category. We believe that the reason is that the January 1993 SIPP values tend to incorporate the SMI premium, while the other 3 SIPP observation points tend to reflect OASDI benefits *net* of the SMI premium. This interpretation is supported by two observations: (1) the usual monthly SMI premium was $36.60 in 1993; (2) the SIPP questionnaire layout was modified after the first 1993 wave to include the bold message, "**Please answer by giving the total amount each month AFTER any deductions such as Medicare premiums**." We note that the conceptual shift from a question about gross OASDI benefits to one on net OASDI benefits occurred in the first wave of the 1993 panel. The mean error for SSI was very close to zero for 2 of the 4 observation points. However, the mean *absolute error* estimate is much larger than the *mean error* for SSI. This is explained by the fact that larger errors in the positive and negative direction largely cancel each other when one looks at mean $ error, but these errors do affect the mean absolute difference.

Table 4 provides a different perspective on estimated SIPP errors. Here we look at unweighted mean SIPP errors as a function of the MBP. We present both mean $ errors and percent errors, because they represent complementary perspectives: some sources of error are independent of the true values (e.g. the SMI premium is a constant independent of the level of benefits), while others may be roughly proportional (e.g. people may fail to report small changes in their income, but more likely to report large percent changes). As Table 4 shows, errors tend to turn from average overestimates to underestimates as the true benefit amount increases. It is not surprising that the percent errors tend to be relatively large at the low end of the distribution of true benefits given that even small $ amounts can be proportionally high at the lowest end of the distribution (e.g. the SMI premium as a percentage of true benefit amounts is inversely related to the level of

Table 4: Discrepancies in January 1993 benefit amount by status in August 1995 sample for adults whose SIPP data are matched to administrative records

	January 1993		August 1995		March 1996		October 1998	
MBP ($)	Error ($)	Error as % of MBP	Error ($)	Error as % of MBP	Error ($)	Error as % of MBP	Error ($)	Error as % of MBP
zero	8	NA	9	NA	7	NA	8	NA
1-100	39	64.2	39	58.4	32	54.7	48	77.1
101-200	57	35.7	23	14.3	20	12.2	28	16.7
201-300	35	13.7	17	6.7	19	7.5	0	0.1
301-400	31	8.9	-11	-3.0	-8	-2.2	-17	-4.9
401-500	14	3.2	-17	-3.9	-28	-6.1	-24	-5.3
501-600	10	1.8	-26	-4.7	-40	-7.2	-56	-10.1
601-700	15	2.4	-35	-5.4	-58	-8.8	-39	-6.0
701-800	-4	-0.5	-43	-5.7	-49	-6.5	-51	-6.8
801-1000	-28	-3.1	-70	-7.9	-75	-8.5	-62	-7.0
1,001-1,500	-97	-8.7	-154	-13.7	-134	-11.8	-133	-11.7
1,501 and over	-52	-3.1	-190	-11.0	-364	-21.2	-340	-20.5
Monthly total SSR benefit ($)	Error ($)	Error as % of SSR	Error ($)	Error as % of SSR	Error ($)	Error as % of SSR	Error ($)	Error as % of SSR
zero	2	NA	3	NA	2	NA	2	NA
1-100	3	5.4	19	38.5	18	35.5	11	21.9
101-200	-14	-9.2	-6	-4.4	10	6.7	15	10.2
201-300	-51	-20.5	-27	-11.3	-45	-18.8	-19	-7.7
301-400	-79	-22.8	-77	-22.9	-48	-14.3	-65	-18.5
401-500	-58	-13.1	-76	-16.6	-83	-17.7	-95	-19.5
501+	-339	-43.6	-169	-26.2	-331	-41.5	-405	-48.8

benefits). What is more remarkable is that percent errors and sometimes even the $ errors tend to be much larger for SSI than for OASDI, particularly at the high end of the distribution. This suggests that the SSI results may be driven by relatively large errors for a small subset of individuals. Occasional large errors may result from topcoding in the 1993 SIPP panel. While the SSR shows individuals in the 1993 SIPP receiving as much as $17,000 from SSI in a single month, the survey values never exceed $1000[4]. The association of true benefit amounts and errors is weaker for OASDI than it is for SSI. However, the magnitude of SIPP underestimate among those in the highest OASDI benefit category ($1501 and over) is substantial for August 1995, and especially for the two 1996 panel observations.

V Other measurement issues

The SSA records provide an outstanding opportunity for analyzing issues related to attrition, refusals, proxy respondents and imputations, because the *administrative records* contain highly accurate information on a small number of characteristics of individuals, and they contain *longitudinal data* on certain outcomes such as monthly benefits, covered earnings and deaths for the full universe of cases with a valid match *regardless of interview status*, time in sample, and other data available for imputations. Due to space limitations we can present only a few preliminary findings in these important areas here.

Table 5: Discrepancies in benefit amounts by type of interview for adults with matched SIPP records

January 1993 (Mean $)	In August 1995 Sample	Not in August 1995 Sample
SSA OASDI (MBP)	113.97	137.01
OASDI error	7.50	6.67
Absolute OASDI error	21.78	28.10
SSA total SSI	6.78	9.11
SSI error	0.39	-0.35
Absolute SSI error	3.49	5.77
N	25,974	6,536

First we look at unweighted data on *survey attrition* (Table 5). We divide the January 1993 sample universe into two groups: those who are also included in the August 1995 sample and those who are not. Since we have administrative record data for both observation points (even for those not in the sample in August 1995), we can compare true benefit amount longitudinally. Leavers tend to have higher average benefits in January 1993 than stayers based on both SSA and SIPP data, but SSA data (not shown) also reveal that this relationship reverses by August 1995. We consistently find that the unconditional (per sample observation) average benefit amounts go up for stayers and go down for leavers. This indicates that attrition is associated with the dynamics of SSA programs. Next we estimate the mean survey error for the two groups for January 1993. We find substantial differences and different patterns for OASDI and SSI. Absolute errors are consistently higher for leavers.

[4] Large payments in the SSR may reflect corrections for earlier underpayments. These large values are also observable in the 1996 SIPP panel, where SSI amounts are not topcoded.

Our next topic deals with the effect of factors associated with *interview status* on survey errors (Table 6). We look at the role of interview status at our first observation point for both panels. We examine the administrative record data both for these two cross-sections and 32 months later. We find that self-reporting interview subjects tend to display a higher level of unweighted unconditional mean OASDI benefit than others, but see no clear patterns with respect to SSI. For the January 1993 sample SSA data reveal a pattern of increasing mean true OASDI benefits; the SSI patterns are less straightforward again. There is substantial variation in mean SIPP errors. The same is true for mean absolute errors. Importantly, mean absolute errors tend to be extremely large relative to the mean of true values for SSI.

Table 6: Discrepancies in benefit amounts by type of interview for adults with matched SIPP records

Interview status	N	*OASDI (net of SMI premium)*				*SSI*			
		MBP ($)	Error	Abs. Error	MBP ($)	SSR ($)	Error	Abs. Error	SSR ($)
1993 panel		January 1993			Aug 95	January 1993			Aug 95
Interview: self	23,180	138.59	9.52	25.41	154.19	7.14	0.86	4.02	7.64
Interview: proxy	8,970	69.07	1.54	15.78	79.85	7.31	-1.20	3.10	8.10
Type Z refusal	128	78.94	-9.97	42.06	85.00	10.00	0.30	20.30	9.95
Type Z other	232	58.01	22.03	57.63	60.23	13.71	-5.39	20.94	11.17
1996 panel		March 1996			Oct 98	March 1996			Oct 98
Interview: self	31,672	145.91	-4.00	26.43	164.84	11.50	-0.57	5.06	11.41
Interview: proxy	10,727	76.14	-4.63	19.26	85.61	9.39	-0.48	5.11	9.35
Non-interview: Type Z	372	66.20	11.15	69.13	79.29	9.70	5.72	25.13	13.78

Next we compare benefit dynamics and survey measurement error by *imputation* status using unweighted data for the first observation point for both panels (Table 7). True mean benefits are consistently high for observations with imputed benefits indicating substantial selectivity. The SIPP error results are dramatic: both mean errors and average absolute errors tend to be much higher for imputed benefits than for survey-reported benefits. The results are particularly striking for SSI. These data suggest that SSA administrative records have substantial potential for reducing SIPP imputation error. The mean errors for nonimputed benefits are very small providing a positive piece of evidence concerning the overall quality of survey reporting, although the relative absolute errors are nontrivial, especially for SSI.

Table 7: Discrepancies in benefit amounts by SIPP imputation status for adults with matched SIPP records

1993 SIPP panel	OASDI (net of SMI premium)				SSI			
	N	MBP ($)	Error	Abs. Error	N	SSR ($)	Error	Abs. Error
January 1993 sample								
Imputed receipt or amount	1,286	345.09	20.55	126.12	102	16.77	29.29	30.75
No imputation	31,224	109.27	6.79	18.80	32,048	7.22	0.15	3.87
August 1995 sample								
Imputed receipt or amount	1,641	435.33	-11.66	119.70	226	19.08	21.82	30.23
No imputation	27,110	113.79	-0.88	18.03	28,525	7.48	1.36	4.39

Finally, we shift our attention from survey measurement error to the effect of the lack of valid SSN match for a subset of SIPP observations using unweighted data. We ask whether the lack of matching SSN's reflects selectivity related to OASDI and SSI benefits. This obviously affects the generalizability of findings from the SSN matched subsample of the SIPP, and has implications on the optimal way to combine survey and administrative record information. (Table 8). Here we can rely only on survey information for comparisons. We find that those without a valid SSN match receive a slightly lower unconditional mean amount of OASDI benefits than those with a valid SSN match. The direction of differences is similar for SSI benefits, but the contrast between the two groups is much more marked in relative terms. These differences merit further investigation since they suggest a pattern of selectivity that may have important implications for combining survey and administrative data for various analytic purposes.

Table 8: SIPP benefit amounts by SER match status for adults in 1993 SIPP panel

1993 SIPP Panel	N	OASDI		SSI	
		Mean ($)	Std. Dev. ($)	Mean ($)	Std. Dev. ($)
January 1993 sample					
Matched with SER	32,510	125.95	278.36	7.49	56.86
Not matched	4,794	113.37	261.74	4.86	45.52
August 1995 sample					
Matched with SER	28,751	130.65	283.76	9.10	64.15
Not matched	4,583	119.09	272.62	7.03	57.02

References

Pedace, Roberto and Nancy Bates (2001), "Using Administrative Data to Assess Earnings Reporting Error in the Survey of Income and Program Participation," *Journal of Economic and Social Measurement* (forthcoming).

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

Pickett, Clark D. and Charles G. Scott (1996), "Reinventing SSI Statistics: SSA's New Longitudinal File," *Social Security Bulletin* 59(2):31-56.

Rupp, Kalman and Paul S. Davies (2000), "Disability Research Using Survey of Income and Program Participation Data Matched to Social Security Administration Record Data," In: *1999 Federal Committee on Statistical Methodology Research Conference: Complete Proceedings*, Statistical Policy Working Paper 30: 605-614, Washington, D.C.: Statistical Policy Office, Office of Information and Regulatory Affairs, Office of Management and Budget.

Rupp, Kalman, Dianne Driessen, Robert Kornfeld and Michelle Wood (1999), "The development of the Project NetWork administrative records data base for policy evaluation," *Social Security Bulletin* 62(2): 1-13.