Reported Earnings in the Survey of Income and Program Participation: Building an Instrument to Target Those Likely to Misreport[1]

Nancy Bates and Robert Pedace
Nancy Bates, U.S. Census Bureau, Room 3376-3, Washington, DC 20023

Key Words: Wage and Salary, Probes

Income Measurement Error in SIPP

The SIPP is a federal survey conducted by the U.S. Census Bureau designed to collect data on personal income sources and amounts, labor force participation, and government assistance program participation and eligibility. One of the uses of these data is to provide statistics on the economic status and distribution of income necessary to address critical public policy questions. One source of determining the accuracy of the SIPP income reports is to benchmark survey estimates to independent aggregate estimates such as the National Income and Products Accounts (NIPA), Social Security Administration (SSA) records, and tax return data. Some studies using these benchmarks suggest that nationally weighted SIPP survey income estimates are consistently below income amounts from benchmark aggregates.

Moore, Stinson and Welniak (1997) present a review of income measurement error studies and make several observations regarding income measurement in the SIPP. First, the quality of income recipiency and amounts in SIPP appears to be improved over the Census Bureau's other major income survey, the Current Population Survey (Coder and Scoon-Rogers, 1996). One notable exception, however, is wage and salary income amounts where the ratio of SIPP aggregate estimates is below both the CPS and independent aggregate estimates.

The tendency for SIPP amount underreporting is also noted in a study comparing wage and salary estimates between the NIPA, CPS and SIPP. Here, Roemer (1999) found the SIPP amount aggregates below both the NIPA and CPS for years 1990-1996. However, both Roemer (2000) and Coder (1992) report only marginal, if any, underreporting in the SIPP respondents' reporting of the *receipt* of wages or salary. In summary, some research suggests that errors in reporting wage and salary *recipiency* is not a major problem in SIPP but that errors in misreporting accurate wage and salary *amounts* present a greater concern.

In this paper, several topics are covered. First, we discuss some methodological alternatives commonly used to address income reporting error. Next, the possibility of programming an automated instrument that would implement reporting improvement techniques (income probes) only among targeted subpopulations is explored. Finally, preliminary analysis from matched administrative record data that allows us to predict which respondents are most likely to misreport wage and salary income recipiency is presented.

Methods to Improve Income Reporting

There are several theories underlying the cause of wage and salary under-reporting in SIPP. Coder and Scoon-Rogers (1996) speculate the shortfall may be attributed to the fact that the SIPP is more conducive to reporting of monthly take home pay rather than gross pay. Additionally, the SIPP may tend to miss 'extra' paycheck amounts from respondents paid on a weekly or biweekly basis. Roemer (2000) hypothesizes another cause could be the disproportionate omission of high-income respondents in SIPP.

When it is suspected that respondents are omitting or underreporting a relatively rare event, one technique commonly used is to determine if a set of probes can be added to aid recall and improve reporting. This can be achieved by adding anything from a simple follow-up probe (e.g., anything else?) to a more detailed series of probes that use examples of atypical situations respondents often fail to recall (e.g., have you received any bonus income? any income from a third payday this month?).

While the addition of probes can be successful in reducing erroneous omissions, they are sometimes problematic for interviewers to implement because they apply to a very small portion of the larger respondent universe. Evidence from behavior coding research suggests that lengthy probes are frequently skipped by interviewers because they are perceived as redundant since they seldom elicit a positive response. The addition of new questions also increases respondent burden and lengthens the interview. If we believe that increased burden

[1] This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a more limited review than official Census Bureau publications. This report is released to inform interested parties of research and to encourage discussion.

contributes to nonresponse, then lengthening the questionnaire to accommodate new probes could be problematic.

CAPI and Programmable Probes

Computer-assisted interviewing (CAI) has become commonplace in both federal and private survey organizations. The SIPP is no exception with the conversion to computer-assisted personal interviewing (CAPI) having taken place in 1996. As a result of the automation, the instrument is reported to better handle complex skip patterns, question fills, and precise tailoring of questions phrasing (U.S. Census Bureau, Demographic Surveys Division; 1998). However, there are additional capabilities offered by automation. For example, the Statistics Canada longitudinal Survey of Labor and Income Dynamic (SLID) uses dependent interviewing to improve recall of income amount reporting based on flags set from previous interviews. If an expected income source is not reported during a subsequent interview, a probe is displayed. This technique increased reporting of certain payment amounts by at least 20 percent (Dibbs, et al., 1995). The SIPP currently uses a similar dependent interview procedure when income receipt is reported during a prior wave by asking during the subsequent wave: "Last time you told me you got income source [type of income]. Do you still receive it?" The SIPP questionnaire designers have recommended making use of additional dependent interviewing techniques in future SIPP panels (Doyle, Martin, and Moore, 2000).

Similarly, it is desirable to have a SIPP instrument that could be programmed to administer probing questions only to that subset of the population for whom they are likely to improve reporting rather than imposing a battery of questions to all wage earners. This would require an imbedded model within the automated instrument that would predict whether a respondent is likely to fail to report income sources and/or misreport income amounts and, if so, administer a set of probing questions only to those respondents.

The first step toward developing such a model is to determine if individuals who misreport earning amounts are easily distinguished from those who do not. The applicability of such a model is contingent upon being able to use predictive information collected in the SIPP prior to the administration of the questions on earnings. Such information includes demographic and labor force characteristics, income recipiency by source of income, and prior wave responses (if not the first wave of interviewing).

The SER Matched to the 1992 SIPP

In order to examine the characteristics of individuals who fail to report income sources or misreport amounts, the 1992 SIPP longitudinal file matched to the Social Security Administration's Summary Earnings Record (SER) is analyzed. The SER contains earnings data that are subject to Social Security taxes. Matching the two files allows for an identification of the occurrence of SIPP income recipiency error and amount misreporting at the micro level using administrative income records linked one-to-one with the SIPP respondents. This represents a unique opportunity to explore person-level characteristics of individuals with income recipiency errors and amount misreporting which is unavailable when using aggregate-level income recipiency and amount data to assess income data quality. For our purposes, the SER income amounts and recipiency reports are considered the 'truth' or gold standard against the SIPP reports. Recipiency error refers to discrepancies in reporting receipt of an income source while misreporting refers to income amount discrepancies (under or overreporting).

The analysis is restricted to respondents that are aged 15 or older and have been assigned a SIPP weight for the calendar year 1992. We constructed an annual earnings variable comprised of SIPP-reported wage/salary and self-employment income. We combined both income sources because the SER contains an aggregate Federal Insurance Contribution Act (FICA) annual earnings amount that includes self-employment income.

Four categories were created based upon the match outcome: (1) zero SIPP and SER earnings; (2) positive earnings recorded on the SER and zero reported on SIPP; (3) positive earnings recorded on both the SER and SIPP; and (4) positive SIPP-reported earnings and zero SER earnings. Table 1 shows the number and percent of individuals in each category.

Empirical Analysis

The analysis focuses on individuals that fall into categories (2) and (3) mentioned above. In particular, the objective of the analysis is to address two issues. First, explore the extent to which earnings recipiency errors are occurring and determine what individual characteristics are associated with this outcome. Second, for those who have earnings on both the SIPP and SER, the goal is to analyze the degree of misreporting (i.e., the difference in the amounts reported in the SIPP and the matched SER) and how this differs among various income categories. In both cases, we are ultimately interested in understanding how

observable individual characteristics can be used to target CAPI income probes and improve response accuracy in the SIPP.

Recipiency Errors

Approximately 3.6 percent of the observations indicate positive earnings on the SER without any SIPP-reported earnings. The median annual earnings for this group based on the administrative data is approximately $925. However, the mean annual earnings for this group is $3,425. These may not appear to be large annual sums, but since these are individuals that have reported no earnings on the SIPP, the expectation is that actual income (from the SER) would also be near zero. On the contrary, some individuals may be failing to report a significant work event (e.g., part-time job, summer job, etc.). In fact, the average recipiency error is approximately 19 percent of SIPP mean total earnings.[2] While relatively few people are in this recipiency error category, the associated amounts are not insignificant.

Next, we examined whether a select number of SIPP demographic characteristics could be used as meaningful predictors of recipiency error. Our response variable is the outcome category in which no earnings are reported in the SIPP, but the SER reports positive earnings (i.e., outcome 2). The following logit model is estimated:

$$L_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad (1)$$

where

$$L_i = \ln\left(\frac{p_i}{1-p_i}\right) \quad (2)$$

the i subscript denotes individual, p is the probability of earnings recipiency error, and X is a vector of demographic characteristics that includes indicators for sec, age, race, ethnicity, marital status, educational attainment, and region of residence.

Table 2 presents parameter estimates and summary statistics for the earnings recipiency error logit model. The covariates that are significantly correlated with earnings recipiency error include some indicators for age, sex, race, ethnicity, and marital status. Region of residence, on the other hand, has no significant correlation with the dependent variable, keeping other factors in the model constant.

Those that appear to be significantly less inclined to recipiency error include individuals aged 30-49 (relative to those aged 15-29), males, and those that are or have been married (compared to those that have never been married). In particular, persons in the 30-49 year old category appear to be almost half as likely to fall into this category relative to the comparison group (i.e., those aged 15-29). The odds of married individuals committing this type of recipiency error are about 41% lower compared to those never married. Similarly, divorced/separated respondents are also less disposed to recipiency error relative to those that have never been married (around 46% less likely).

On the other hand, individuals aged 50-64, blacks, Hispanics, and those with less than a college education appear to be significantly more inclined toward recipiency error relative to each respective comparison category. Individuals aged 50-64 and those with less than a college education are each approximately 30 to 38 percent more likely to commit recipiency error compared to those aged 15-29 and persons with some college education, respectively. The most striking recipiency error probabilities, however, are those of blacks and Hispanics who are approximately 70 and 50 percent more disposed to this type of recipiency error compared to whites and non-Hispanics, respectively.

Misreporting

In order to accurately compare earnings reported on the SIPP to those on the SER, the SIPP amounts were capped corresponding to the Old-Age, Survivors, and Disability Insurance (OASDI) maximum taxable amount. According to the Federal Insurance Contributions Act (FICA) employers, employees, and the self-employed are required to pay taxes on earnings in covered employment up to an annual maximum taxable amount for OASDI (Social Security Administration, 1998). The OASDI maximum is updated each year in proportion to increases in nationwide average wage and salary earnings (Social Security Administration, 1998). In 1992 the OASDI maximum taxable amount was $55,500.

One way to explore the misreporting issue is to examine the characteristics of the SIPP and SER dollar amount discrepancies for different income categories. Figure 1 displays the fraction of SIPP respondents that report earnings receipt that is within a given percentage of the SER amount.

The magnitude of misreporting clearly varies by income category. In the lowest earnings

[2] SIPP capped mean annual earnings is $18,102. A 'cap' was placed on SIPP income amounts to make them comparable to SER amounts -- the SER maximum for 1992 was $55,500.

category (i.e., those with SER amounts of less than $5,000), less than 40 percent have SIPP reported earnings within 25 percent of the SER amount. As earnings increase, however, the magnitude of misreporting, as a proportion of SER amounts, declines. For those earning between $20,000 and $40,000, approximately 80 percent report SIPP earnings within 25 percent of the SER amount. At this point the trend appears to reach a plateau. The earnings cap limitations on the SER data, however, do not allow us to accurately examine misreporting for the highest income categories.

Another interesting characteristic of SIPP misreporting is that, on average, it changes from overreporting in the lowest earnings categories to underreporting for those with at least $20,000 in annual earnings (see Figure 2). Two possible explanations of overreporting by individuals in the lowest income categories is that they may hold jobs that are not covered by the OASDI program or they are simply young workers receiving unreported income.

The expectation, therefore, is that individuals in these categories would be, predominantly, in government and farming/agricultural jobs. However, the industry, occupation, and worker class distributions for this group does not reveal a pattern of employment in these uncovered sectors. On the contrary, workers appear to be dispersed over a wide range of industries and occupations. Also, the median age for those earning less than $5,000 and between $5,000 and $10,000 is 27 and 33 (with means of 32 and 36), respectively. This is not suggestive of youth employment in the informal sector.

The issue of over- and underreporting, and its magnitude must be investigated further in order to understand how the CAPI instrument can be modified to improve SIPP response accuracy. One obvious extension is to build models that predict type and degree of income amount misreporting in addition to income recipiency error.

Conclusions/Future Research

A preliminary examination of our matched file suggests that the reporting error of wage and salary recipiency omission is the rarest possible outcome in the SIPP (3.6% of matched cases indicated income recipiency according to SSA records, but no reported wage and salary income according to the SIPP). This is an encouraging finding and supports earlier studies indicating that the number of wage and salary recipients measured in the SIPP is very close to the 'truth.'

What are the characteristics of respondents that should be targeted? At this stage in the analysis the predictors were restricted to only a few basic demographics. However, our preliminary model indicates that those aged 50-64 are more likely to omit income receipt compared to the youngest age group (15-29). The odds of omitting income recipiency for men are lower compared to women and compared to whites, blacks were found significantly more likely to exhibit recipiency error. Respondents reporting a Hispanic origin are more likely to omit income receipt compared to non-Hispanics and all three education categories below 'some college' appear more likely to omit income receipt. Finally, respondents who were currently married (or had been at some time, i.e., widowed, divorced/separated) were significantly less likely to omit receipt of income compared to respondents who were never married. It was also discovered that as respondents' income increases, the amount of income reporting discrepancy tends to decrease as a fraction of total earnings.

Lacking any other information, the instrument could be programmed to target income probes only for those respondents with some combination of the above noted characteristics. However, our profile of recipiency error at this stage is probably premature (as evidenced by the poor overall fit of the model). Additionally, there are probably significant interactions between the different characteristics that need to be explored when trying to predict recipiency error. Finally, we need to explore misreporting as our response variable, to learn, for example, the characteristics associated with over and under-reporting income amounts. Consequently, the set of response and predictor variables presented here must be refined before recommending a test of new automated probes.

References

Coder, J. and Scoon-Rogers, L. "Evaluating the Quality of Income Data Collected in the Annual Supplement to the March Current Population Survey and the Survey of Income and Program Participation." *SIPP Working Paper Series*, No. 215, U.S. Bureau of the Census, Washington, D.C. 1996.

Coder J. "Using Administrative Record Information to Evaluate the Quality of the Income Data Collected in the Survey of Income and ProgramParticipation." *Proceedings of Statistics Canada Symposium 92: Design and Analysis of Longitudinal Surveys*, Ottawa: Statistics Canada. 1992.

Dibbs R., Hale, A., Loverock, R., and Michaud, S. "Some Effects of Computer Assisted Interviewing on the Date Quality of the Survey of Labour and Income Dynamics. *Proceedings of the International Conference on Survey*

Measurement and Process Quality. Alexandria, VA: American Statistical Association (1995), 174-177.

Doyle, P., Martin E., and Moore, J. "Methods Panel to Improve Income Measurements in the Survey of Income and Program Participation." *A paper presented at the Conference of the American Association for Public Opinion Research*, Portland , Oregon, (2000).

Moore, J., Stinson, L., and Welniak, E. "Income Measurement Error in Surveys: A Review." A paper presented at the *Cognitive Aspects of Survey Methodology II Conference*, Charlottesville, VA, 1997.

Roemer, M. "Reconciling March CPS Money Income with the National Income and Product Accounts: An Evaluation of CPS Quality." A paper presented at the *Annual Meetings of the American Statistical Association*, Baltimore, MD, 1999.

Roemer, M. "Assessing the Quality of the March CPS and SIPP Income Estimates, 1990-1996." A paper presented at the *American Statistical Association International Conference on Survey Nonresponse*, Portland, OR. 1999.

U.S. Census Bureau, Demographic Surveys Division. "Automated Data Collection: How Have the Demographic Surveys Fared Thus Far?" April, 1998.

U.S. Census Bureau, Current Population Report, P60-193. *Money Income in the United States: 1995 (With Separate Data on Valuation of Noncash Benefits*. U.S. Government Printing Office: Washington, D.C. 1996.

U.S. Department of Commerce, U.S. Census Bureau. SIPP Quality Profile. *SIPP Working Paper* No. 230. 3rd Edition. 1999.

TABLE 1. Cross-Tabulation of SIPP Reported Earnings Receipt by Matched SSA Summary Earnings Records (SER)

Income Receipt Reported in SIPP?	SER Indicates Earnings Receipt?		Total
	No	Yes	
No	9,346	1,150	10,496
(%)	(29.0)	(3.6)	(32.6)
Yes	2,070	19,617	21,687
(%)	(6.4)	(61.0)	(67.4)
Total	11,416	20,767	32,183
	(35.4)	(64.6)	(100.0)

Note- Unweighted figures are presented in order to show the actual sample densities in each category. here are no significant differences between these and the weighted tabulations

TABLE 2. Logit Estimates for Earnings Recipiency Error

Predictor	*Parameter Estimate*	*Odds Ratio*	*Standard Error*
Age (15-29):			
30-49	-0.586***	0.557	0.093
50-64	0.280**	1.323	0.101
65+	0.047	1.049	0.117
Sex (Female):			
Male	-0.248***	0.780	0.062
Race (White):			
Black	0.537***	1.711	0.086
Asian/Pacific Islander	0.222	1.249	0.182
Am. Indian/Eskimo	0.025	1.026	0.365
Ethnicity (non-Hispanic):			
Hispanic	0.391***	1.478	0.099
Education (College):			
No School	0.325*	1.384	0.206
Elementary School	0.261**	1.298	0.125
High School	0.300***	1.350	0.068
Marital Status (Never Married):			
Married	-0.529***	0.589	0.086
Widowed	-1.026***	0.358	0.168

Table 2 (con't.)

Divorced/Separated	-0.610***	0.544	0.129
Region (East):			
South	-0.018	0.982	0.081
Midwest	-0.050	0.951	0.088
West	0.033	1.034	0.095

Sample Size=32,183; Log Likelihood=-4,792.07; Generalized R^2=0.039; $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$

Note- The dependent variable is equal to 1 if the individual reported no SIPP earnings but had positive SER earnings, 0 otherwise. The response profile is such that this variable is equal to 1 for 1,150 individuals and equal to 0 for 31,033 individuals.

Figure 1.

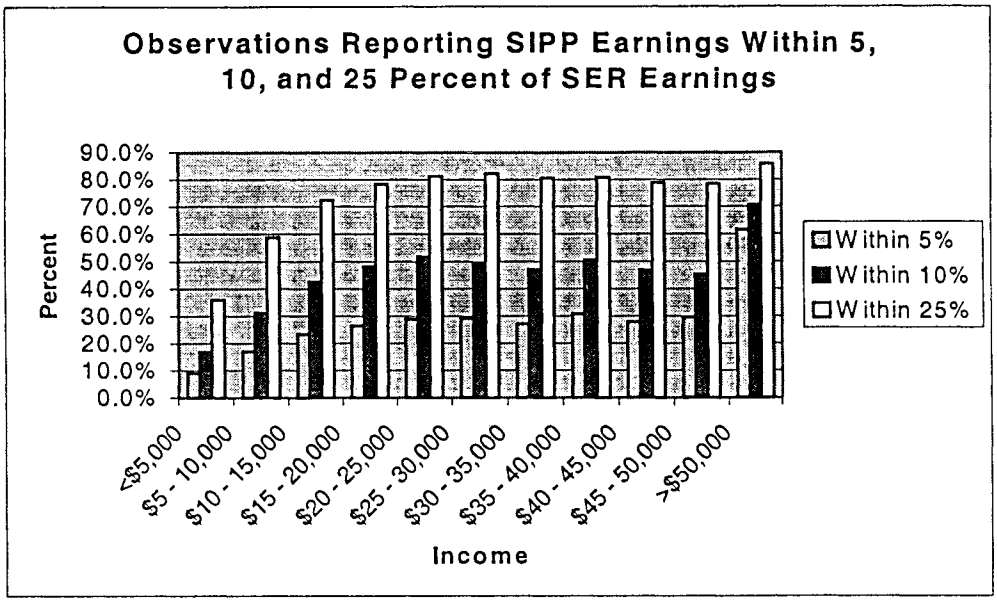

Figure 2.