RESEARCH REPORT SERIES
*(Survey Methodology #2006-4)*

# The Effects of Questionnaire Design Changes on General Income Amount Nonresponse in Waves 1 and 2 of the 2004 SIPP Panel

Jeffrey C. Moore

Statistical Research Division
U.S. Census Bureau
Washington, DC 20233

Report Issued: March 1, 2006

# The Effects of Questionnaire Design Changes on General Income Amount Nonresponse in Waves 1 and 2 of the 2004 SIPP Panel

Jeffrey C. Moore
Statistical Research Division
U.S. Census Bureau

Abstract

The "general income" section of the SIPP interview asks about income other than earnings, and other than financial investments – it focuses primarily on participation in and receipt of income from pensions, public assistance, and other transfer-type programs. New procedures were introduced in the 2004 panel wave 2 SIPP questionnaire in an attempt to reduce nonresponse to general income amount items. The new procedures consist of the ability to use dependent follow-up questions of the following sort in the event of an initial nonresponse to an amount question:

> "It says here that you received $550 from workers compensation last May. Does that still sound about right for June and July?"

This report summarizes the impact of these new procedures on nonresponse to amount items for selected general income sources in wave 2 of the 2004 SIPP panel, through a comparison with the same estimates from the preceding (2001) SIPP panel. The paper also examines nonresponse to wave 1 amount items for the same income sources. Major findings are as follows:

(1) The new dependent nonresponse follow-up procedures had a consistent and major positive impact on final nonresponse in 2004 wave 2.
(2) On the negative side, however, is evidence of improper use of the dependent follow-up procedures by SIPP interviewers, resulting in very high rates of initial nonresponse to the wave 2 amount items in the 2004 panel.
(3) Despite the fact that income reporting procedures in wave 1 were essentially identical across the two panels, there also appears to have been a consistent decline in wave 1 nonresponse in 2004 compared to 2001. Although the evidence is far from definitive, it appears that a new introduction to the income questions may have helped allay privacy concerns, and thus reduced refusal nonresponse.

Keywords: dependent interviewing, income amount item nonresponse, interviewer behavior, questionnaire design

**The Effects of Questionnaire Design Changes on General Income Amount Nonresponse in Waves 1 and 2 of the 2004 SIPP Panel**

Jeffrey C. Moore
Statistical Research Division
U.S. Census Bureau

Executive Summary

New dependent follow-up procedures were introduced in the wave 2 "general income" section of the 2004 panel SIPP questionnaire, in an attempt to reduce nonresponse to questions about income received from pensions, public assistance, and other transfer-type programs. In the event of an initial nonresponse to an amount question, a dependent follow-up question of the following sort was asked:

> "It says here that you received $1,850 from your company or union pension last March. Does that still sound about right for the last four months?"

This report summarizes the impact of the new follow-up procedures on nonresponse to income amount items for several selected general income sources in wave 2 of the 2004 SIPP panel, through a comparison with the same estimates from the preceding (2001) SIPP panel. The paper also examines nonresponse to wave 1 amount items for the same income sources. Major findings are as follows:

(1) The new dependent nonresponse follow-up procedures had a major positive impact on final nonresponse in 2004 wave 2. For all 10 general income sources examined, final nonresponse in wave 2 (after implementation of the follow-ups) was significantly lower in 2004 than it was in 2001. In addition, while the consistent trend in the 2001 panel results was for item nonresponse rates to increase significantly from wave 1 to wave 2, the addition of the follow-up procedures produced the opposite trend in 2004, where final wave 2 item nonresponse rates were consistently *lower* than the wave 1 rates.

(2) On the negative side, however, is evidence of improper use of the dependent follow-up procedures by SIPP interviewers, resulting in very high rates of initial nonresponse to the wave 2 amount items in the 2004 panel.

(3) There appears to have been a general decline in wave 1 nonresponse in 2004 compared to 2001 as well. This effect is of smaller magnitude than the typical wave 2 difference, and is not consistently significant across all sources examined. Nevertheless, its appearance in the results is surprising, since the income reporting procedures in wave 1 were essentially identical across the two panels. A possible explanation may reside in the new introduction to the income questions in the 2004 questionnaire, which was intended to allay privacy-related reluctance to reveal income amounts, and may have done so.

## 1. Introduction and Overview

Income data are among the most difficult data to collect in surveys, and as a result often suffer from high rates of item nonresponse (Moore, Stinson, and Welniak, 2000); the U.S. Census Bureau's Survey of Income and Program Participation (SIPP) is by no means immune to this problem. A new SIPP core questionnaire was introduced at the start of interviewing for the 2004 SIPP panel, following several years of development and testing through the SIPP Methods Panel project (Doyle, Martin, and Moore, 2000; Moore, Pascale, Doyle, Chan, and Griffiths, 2004). Although the primary goal of the Methods Panel project was to reduce unnecessary burden and improve the SIPP interview process, some of the new procedures focused on income amount item nonresponse as well. This paper examines the impact of those procedures on 2004 panel nonresponse to income amount items for "general income" sources – i.e., pensions, public assistance, and other transfer-type programs.

The new procedures consisted of the capability, after the initial interview wave, to use a dependent question as a follow-up in the event of an initial nonresponse to a general income amount item – "I have recorded from last time that you received [dollars] from [income source] last February. Does that still sound about right for March, April, and May?" No such dependent interviewing procedures were used in earlier SIPP panels.

This paper evaluates the effectiveness of the new procedures at reducing nonresponse to amount questions for general income sources, by comparing nonresponse levels in the new panel to those of the most recent panel before the new procedures were introduced, the 2001 SIPP panel. The remainder of this paper is organized as follows: Section 2 provides a brief description of the SIPP survey, and the research and development program which led to a new core questionnaire in the 2004 SIPP panel. Section 3 describes in some detail the new procedures introduced into the 2004 SIPP questionnaire in an attempt to reduce nonresponse to general income amount items. Section 4 summarizes the analysis procedures used, and presents and discusses their results – noting, of course, the limitations of the research. Finally, section 5 draws some conclusions concerning the implications of this research, as well as directions for future research.

## 2. SIPP Background

### 2.1. Basic SIPP Design Features

SIPP is a nationally-representative, interviewer-administered, longitudinal survey conducted by the U.S. Census Bureau. SIPP data are used to analyze and evaluate income, wealth, and poverty in the United States, the dynamics of program participation, and the effects of government programs on families and individuals. The current (2004) SIPP panel is expected to consist of fifteen waves (or rounds) of interviewing over five years, with waves administered three times a year, at four month intervals. (Historically, panel length has varied from less than a year to up to four years, depending on such factors as budget constraints, redesign schedules, coordination

with decennial census activities, etc.). The SIPP sample is split into four equivalent subsamples, or "rotation groups;" each rotation group's interview schedule is staggered by one month, in order to maintain a constant workload for field staff. Since 1996, all SIPP interviews are conducted with a computer-assisted questionnaire; the first interview is administered in-person, subsequent interviews are often conducted via telephone. The SIPP core instrument, which contains the survey content that is repeated in every survey wave, is detailed, long, and complex, collecting information about household structure, labor force participation, income sources and amounts, educational attainment, school enrollment, and health insurance over the prior four-month period. A wide variety of other specific topics are addressed in the survey's "topical modules," which are appended to the core interview and which may appear once per panel, or biennially, or annually. A typical SIPP interview takes about 30 minutes per interviewed adult. See U.S. Census Bureau (2001) for a more complete description of the SIPP program.

### 2.2. Testing and Refining New Questionnaire Procedures in the SIPP Methods Panel

Motivated primarily by concerns about increasing unit nonresponse and attrition, and by a desire to bring known data quality problems under better control, in the mid-1990's the Census Bureau launched a research and development program to create an improved SIPP questionnaire for implementation in the new SIPP panel to begin in 2004. The SIPP Methods Panel Project was created to design improved procedures and to test and refine them in a series of split-panel field tests. For the most part, the Methods Panel's efforts focused on "interview process" improvements that would yield a less burdensome interview. An important secondary focus, however, was improved data quality in selected areas. One particular such improvement goal was a reduction in nonresponse to income amount items, especially for asset income sources, where nonresponse rates as high as 40 to 50 percent are not uncommon (Moore, Stinson, and Welniak, 2000). The new procedures also targeted nonresponse to general income amount questions, although less intensively, since the nonresponse problem for general income sources is not nearly so severe as it is for assets.

The Methods Panel project included a series of three field experiments to evaluate and refine an experimental SIPP questionnaire in preparation for its implementation as the production instrument in the 2004 panel. The first of these tests was carried out in the late summer of 2000 (MP2000), and included only a wave 1 interview. The second two field tests – MP2001 and MP2002, carried out in the summer and fall of 2001 and 2002, respectively – included both a wave 1 and a wave 2 interview. The Methods Panel field experiments employed an experimental design, with random assignment of sample cases to test and control groups. Sample households assigned to the control treatment were administered the standard 2000/2001 panel SIPP questionnaire; those in the test treatment were interviewed with the experimental instrument. Interviewers' assignments included a mix of both the control and experimental instruments, so they could provide feedback on the relative merits of each, and also to avoid confounding staffing differences with instrument treatments. The sample size for each test was set at a level which was expected to yield about 2000 completed wave 1 interviews (i.e., 1000 per treatment). Based on post-test evaluations, adjustments and refinements were made to the experimental

questionnaire after each field test. (See Doyle, Martin, and Moore (2000) for a more complete description of the design of the field experiments; Moore, et al.(2004) also covers field test design issues in some detail, and provides more general information concerning the changes implemented in the SIPP questionnaire.)

The results of the field tests with regard to item nonresponse for general income amounts were sufficiently positive (Moore and Griffiths, 2003) that the new procedures were implemented in the redesigned production instrument developed for the 2004 SIPP panel. The next section describes those new procedures in some detail.

## 3. New Procedures to Reduce Nonresponse for General Income Amounts

Two of the new procedures employed to reduce income amount nonresponse for asset income sources – the use of a flexible reporting period, and a closed-ended "range" nonresponse follow-up question (see Moore, 2006) – were not incorporated into the general income section of the questionnaire. First, the SIPP questionnaire has always asked respondents to report monthly income amounts from general income sources, and monthly income is precisely what many (if not most) of those sources actually supply. Therefore, it was not deemed necessary to further complicate the questionnaire by offering reporting period flexibility, since it appeared that the most convenient reporting period for most general income sources was already in place. Second, while an argument could be made that the addition of some sort of "range" follow-up questions might have been useful, other factors argued against their inclusion. Most importantly, unlike the situation for assets, these procedures – and especially the processing software to make effective use of them – were not part of the long-standing and familiar "culture" of this section of the SIPP questionnaire, nor were resources readily available to implement the changes that would have been necessary to both the questionnaire and the processing systems. And, as noted, the magnitude of the nonresponse problem for general income sources was not nearly as great as it was for assets, thus minimizing the need for more "radical" solutions. For these reasons, the only new procedures implemented in this section of the questionnaire to combat income amount nonresponse were dependent follow-up questions, beginning in wave 2, very much like those used in the assets section of the questionnaire (Moore, 2006).

As with assets, one noteworthy feature of the new dependent follow-up procedures was the addition of a new response option, "L," to the initial general income amount questions starting in wave 2. This was intended to provide an obvious place to record a very common type of response, according to SIPP interviewers: "nothing has changed," or "it's the same as last time," or other words to that effect. (As we shall see, there is strong evidence that this is not how interviewers used it.) Thus, if the new RIP[1] procedures did not preclude mention of the prior

[1]The 2004 SIPP panel saw the first implementation in SIPP of the Respondent Identification Policy (RIP). With the implementation of this policy, survey information reported by one member of a household cannot be shared with other household members without the consent of the original respondent.

wave's report in the current wave, a new response option, the "L" option, was displayed on the interviewer's computer screen as a possible response to the initial amount question:

> FSAMT15 (2004)
> How much has John received in Food Stamps ...
> ENTER (N) FOR NONE OR NO MORE
> ENTER (L) FOR SAME AS LAST INTERVIEW
>
> ... so far this month? $_____
> How about in September? $_____
> And August? $_____
> And July? $_____
> And June? $_____

Starting in wave 2 of the 2004 panel, nonresponse to an initial general income amount item of the sort shown above – that is, an L entry for the first month asked about, or all D's or R's for all months – generally elicits a dependent question such as the following:

> PWFS (2004)
> It says here that you received *[previous wave monthly dollar amount]* in food stamps benefits last *[month of previous wave amount]*. Does that still sound about right for the last four months?

A "no" response to the dependent follow-up elicits the following correction screen:

> PWFSFIX (2004)
> Income Type: FOOD STAMPS
> What is the correct monthly amount?

As noted, the dependent question is *generally* elicited following an initial nonresponse. There are circumstances when use of this option is not feasible – such as when the income source in question has only recently begun to provide income, or when the respondent was not interviewed in the preceding wave, or failed to provide amount information in the prior wave, or when the prior wave respondent did not consent to the "RIP" request to allow his/her answers to be revealed to another member of the household. In such circumstances the 2004 questionnaire has no special procedures if the initial nonresponse was a D or an R. However, if an L was entered and a dependent follow-up question cannot be used, the new questionnaire asks for an approximate amount:

FSSORRY (2004)
I'm sorry. That information doesn't seem to be in my computer. Can you give me an approximate amount?

## 4. Analysis

### 4.1. Analysis Strategy and Construction of Nonresponse Estimates

This section summarizes the evaluation of the 2004 panel instrument changes, described above, on nonresponse to general income amount questions. The general income section covers over 40 different income sources; this analysis selects 10 of those sources for evaluation. The basic evaluation strategy involves a comparison of nonresponse to the amount questions for those specific sources in wave 2 of the 2001 and 2004 panels. The analysis looks at both initial nonresponse in 2004, before implementation of any nonresponse follow-up questions, and final nonresponse, after any nonresponse follow-ups. Although no new procedures were introduced to address nonresponse in wave 1, the evaluation also includes wave 1 nonresponse estimates, largely to permit a very preliminary assessment of the impact of the new nonresponse follow-up procedures on nonresponse trends across the life of a SIPP panel.

With the exception of the new dependent follow-up questions in 2004 wave 2, the 2001 and 2004 panel questionnaires asked about general income amounts in very similar ways. The main difference is that the 2004 questionnaire asks about income received in the current (interview) month and the 2001 questionnaire did not. The evaluation compensates for this difference by simply ignoring interview month information[2]. Thus, the nonresponse estimates were calculated very simply, as follows: For each income source, each respondent was asked 0, 1, 2, 3, or 4 amount questions – one for each month he/she reported receipt of any income from that source – and, similarly, each respondent failed to provide an initial amount report (i.e., had a D, R, or L entry in the amount report field) 0, 1, 2, 3, or 4 times. Taking the wave 2 follow-up questions into account – that is, treating either the confirmation of the dependent, last-wave amount, or the reporting of a better amount, as if it negated all of that respondent's initial nonresponses – each respondent also had a final nonresponse tally of 0, 1, 2, 3, or 4. The number of amount questions asked was summed across all respondents, forming the denominator of the nonresponse estimate; the number of initial nonresponses was summed across all respondents, forming the numerator for the initial nonresponse rate, and similarly the summed number of final nonresponses became the numerator for the final nonresponse rate.

[2]Although data for the current month were ignored in the analysis, in some circumstances the current month report affected the data recorded for other months. The 2004 questionnaire was designed so that once an "L" was entered for any relevant month, the dependent follow-up question appeared immediately, thus skipping over all remaining months' amount questions. The analysis treats the single "L" response as applying to all months in which the respondent reported receipt of income from the source in question. In this way the current month report, if it was an "L," affected the data for other months, even though the current month itself was ignored in the analysis.

4.2. Data Sources

The analysis uses the "TransCASES" data files from the first two waves of the SIPP 2001 and 2004 panels, unweighted. These are "raw" data files derived directly from the survey instrument without any data editing (except for the limited edits performed as part of the operation of the instrument itself) or imputation. Thus, the estimates in this report may not match exactly with those derived from subsequent datasets which have been edited, and which include imputed data.

4.3. Results

Table 1 presents the basic nonresponse estimates, showing for waves 1 and 2 in each panel the overall nonresponse rates (expressed as percents) for 10 different general income sources, across all respondents whose interview included at least one amount question for that source. For 2004 wave 2, Table 1 shows both an initial rate, before implementation of the nonresponse follow-ups, and a final rate. According to the results of simple t-tests, virtually all relevant comparisons show a statistically significant difference at the $p<.10$ level or beyond; matching superscripts indicate the few comparisons for which that is not the case.

Table entries may be read as follows (referring, for example, to the results for Social Security): In wave 1 of 2001 20.0% of the 48,643 questions respondents were asked about a monthly amount of Social Security income were met with a nonresponse – a D (don't know) or R (refused). This rate is significantly higher than the 15.8% rate of nonresponse observed in 2004 wave 1. The nonresponse rate in wave 2 of 2001 increased significantly compared to the wave 1 estimate, rising from 20.0% to 23.4%. Although the initial nonresponse rate in 2004 wave 2 was very high – 56.0% – the final rate, after the administration of the dependent follow-up question, dropped to 11.4%, which is significantly lower than the 2001 wave 2 estimate, and also indicates a significant decline in nonresponse from the level observed in wave 1 of the 2004 panel. The total number of Social Security amount questions asked in wave 1 (i.e., the denominators for the nonresponse estimates) were 48,643 and 62,087 in the 2001 and 2004 panels, respectively, and 39,522 and 59,787 in wave 2.

4.3.1. Impact of nonresponse follow-up procedures on item nonresponse in wave 2

Table 1 provides evidence of a consistent, positive impact of the new dependent follow-up procedures on final rates of item nonresponse for these general income sources. For each of the ten income sources, final nonresponse in 2001 increased significantly from wave 1 to wave 2. In 2004, with addition of the dependent procedures in wave 2, exactly the opposite is true, and all wave 2 nonresponse estimates are lower than they were in wave 1 (only the "State Unemployment" difference is non-significant). Additional evidence of the positive impact of the new procedures can be seen in the large and consistent final nonresponse rate differences within wave 2 – in every case the 2004 estimate is significantly lower than the 2001 rate.

Another unmistakable feature of the results in Table 1 is the extremely high rate of initial nonresponse in wave 2 of the 2004 panel, which is in every case significantly higher than the wave 1 rate, and in every case significantly higher than the 2001 wave 2 rate. The major difference in wave 2 of the 2004 panel, of course, and a very likely factor in the increase in initial nonresponse, is the presence of the dependent nonresponse follow-up questions. Although definitive proof is lacking, several trails of evidence point to interviewer behavior as an important cause of the major jump in initial nonresponse. One such trail is comprised of several 2004 panel interview observation reports from wave 2 and beyond (Bruun, 2005; Davis, 2005; Gilbert, 2005; Moore, 2004), each of which independently notes interviewers' misuse of the dependent nonresponse follow-up option as a means to "peek ahead" at what the respondent reported in the last interview, before he/she even had a chance to report an amount on his/her own. E.g.:

> There were a significant number of questions ... which the FR did not even ask... The FR frequently skipped the question asking how much an individual earned from a given asset. He used the 'L' option and asked the respondent whether or not that value was still the case. [Bruun (2005), p1]

> The FR [said] that he is hesitant to offer the "L" option, but it seemed that he was always prepared to offer it. ... he would always check to see if there was an amount from last time before the respondent could answer. [Gilbert (2005), p2]

> Any time a respondent hesitated even for an instant to provide an answer to an amount question – well before an explicit "I don't know" or "It hasn't changed" issued from the respondent's lips – [the interviewer] voluntarily offered to check to see if there was an amount from last time which could be used as a prompt. [Moore (2004), p3]

The data shown in Table 2 offer strong confirmatory evidence that the high rate of initial nonresponse is a result of interviewer behavior, and not respondent conditioning. This table summarizes initial 2004 wave 2 responses to the income amount questions asked about a broad group of general income sources, all of which use the same questionnaire screens to capture amount information, and two other major transfer programs whose income data are captured separately. Most importantly, the results are shown separately by rotation group, each of which is comprised of a different set of respondents. Note the large upward trend in use of the "L" option across the four rotation groups: in all cases the rate essentially doubles from rotation group 1 to rotation group 4. Obviously, this trend is independent of respondents' actions, and can only reflect changes in interviewer behavior.

There is no mention of this problem in any of the observation reports from the Methods Panel testing phase. However, during recent consultations with SIPP headquarters staff on the development of refresher training materials to try to curb overuse of the "L" option, it became clear that there were some important misperceptions about the "L" option. The Methods Panel intent was that "L" was a response option like any other, to be recorded by the interviewer only if

respondents uttered appropriate words (i.e., words to the effect that the amount in question was "the same as the last interview"). The refresher training development surfaced another interpretation entirely, which was that the presence of the "L" option was a signal to interviewers that a prior wave answer was available for their examination.

This misperception triggered a review of the SIPP wave 2 training materials which found some unintended support for it. Nowhere do the training materials explicitly note that "L" is only to be entered if "same as last time" (or something similar) is what the respondent actually *says*; instead, the impression is left that interviewers may access the prior wave information whenever it is available. Sometimes this permission is implied:

> Throughout each section of the instrument, you will see the use of dependent data. As you know, dependent data is information provided by a respondent in the previous interview. It is only used when:
> – There is no Respondent Identification Policy (RIP) problem;
> – The person is a continuing respondent, that is, the person was a household member during the previous interview, and
> – We have data for that particular question from the previous interview.
> ...
> **If the respondent answered "yes" to the RIP question in the previous interview, we are free to use dependent data in the next interview, no matter who in the household we speak with.** [U.S. Census Bureau (2004), pH-3, emphasis added; also pI-3]

And sometimes it is quite explicit:

> [as part of a practice interview, an interviewer is conducting the "earnings" section of a wave 2 interview, reading the SAME screen:]
> I have recorded from before your biweekly paychecks are the same each time. What is that gross biweekly amount before deductions?
> [note to instructor:] **(Interrupt and say:)**
> **This is an example of dependent data from the last interview. Go ahead and enter "L" now.** [U.S. Census Bureau (2004), pH-93, emphasis added; also pI-70]

Thus, it is possible that interviewers' misuse of the "L" option derives from an initial training that was at best unclear, and at worst misleading. A wave 6 refresher training package was developed in an attempt to address the problem; whether it can effectively counteract a misunderstanding that has gone unchallenged for several survey waves remains to be seen.

4.3.2. Item nonresponse differences in wave 1

Although the main focus of the analysis of item nonresponse for general income amounts has been the new wave 2 procedures, another feature of Table 1 also demands some attention. In addition to the wave 2 nonresponse differences described and discussed above, there appears to

have been a systematic decline in wave 1 nonresponse in 2004 compared to 2001 as well. The effect is of smaller magnitude than the typical wave 2 difference, and is not consistent across all sources examined (i.e., the observed differences for both State Unemployment and Alimony are non-significantly in the "wrong" direction). Nevertheless, its appearance in the results is somewhat unexpected, since the income reporting procedures in wave 1 were essentially identical across the two panels.

The first factors examined in the search for a possible cause of this difference had to do with how the interviews were conducted: was there a reduction in proxy reporting in 2004, or an increase in the use of records, that might explain a reduction in wave 1 nonresponse in 2004? The answer to both questions is no. The rate of proxy reporting in wave 1 of the two panels was identical: 33%. And, although the comparison is muddied somewhat by differences in how the data were collected[3], respondents' use of records to assist reporting of wave 1 general income amounts seems to actually have declined somewhat in 2004. In 2001 interviewers reported that 27.4% of those who answered general income amount questions used records at least once, which is substantially (and significantly) greater than the 22.5% who used any records in 2004.

The only difference between the two panels in how wave 1 income data were collected has been alluded to already – in 2004 amounts received in both the interview month and the preceding four calendar months were captured; 2001 included only the preceding four calendar months. Perhaps the more easily recalled information for the current month served as some sort of catalyst for improved recall in subsequent months, thus reducing nonresponse across the board. This hypothesis, too, fails to find support in the data. First, there is little evidence that amount reporting for the current month is particularly easy relative to the other months. In fact, nonresponse for the first month that the respondent reported about was quite consistent, regardless of whether the first month reported about was the current month, or "last month," or any of the other reference period months (data not shown). Second, if the current month's report somehow facilitated recall of information for subsequent months, then we would expect to see the impact take the form of reduced "don't know" nonresponse in 2004. In fact, the nonresponse reduction in 2004 appears to have been a result of reduced refusals (9.1% for the first amount question asked in 2004, compared to 13.2% in 2001) – "don't know" nonresponse was actually slightly up in 2004 (5.2%) compared to 2001 (4.7% – data not shown)[4].

A close comparison of the 2001 and 2004 questionnaires[5] revealed one additional difference that, if only by process of elimination, may help explain the improvement in nonresponse in 2004.

---

[3]The 2001 questionnaire included a separate question (for the interviewer) about the respondent's use of records for each general income source reported. The 2004 questionnaire, in contrast, asked only a single question at the end of the general income amounts section, covering all general income sources combined.

[4]These results summarize responses to the MNTHAMT15 screens, which captured income amounts for most (but not all) general income sources.

[5]I thank Tim Gilbert (DSD) for finding and reminding me of this subtle change in the 2004 questionnaire.

The SIPP core questionnaire is structured such that it first obtains a listing of all income sources – jobs and businesses, general income sources, and assets – and then focuses on income amounts from those reported sources. As part of the Methods Panel process we tested a change to the introductory text to be read to respondents at the beginning of the income amount section. The introduction included in the 2001 questionnaire placed substantial emphasis on the use of records:

> ADINCSRCE2 (2001)
> The next part of the interview is about those sources of income from October 1st through the end of January. Since accuracy is important to this survey, it would be very helpful if you could refer to any income records you might have for the next series of questions. I would be happy to wait while you get them. Do you need a moment?
> GIVE RESPONDENT A CHANCE TO GET RECORDS.
> PRESS "ENTER" TO CONTINUE

Several factors motivated the exploration of possible improvements to the introduction, including especially interviewers' negative attitudes (they were much put off by its heavy-handed approach to record use), and the absence of any indication that it resulted in increased record use or improved data quality. After testing, a new introduction was incorporated in the 2004 questionnaire, one with a much-softened appeal for record use, and with new, optional text which was intended to allay respondents' privacy-related concerns and otherwise encourage income amount reporting:

> ADINCSRCE2 (2004)
> The next part of the interview is about your income since October 1st. We want to be as accurate and efficient as we can, so it would be very helpful if you could refer to any records you might have. (Also – we know that people aren't used to talking about their income, but we ask these questions to get an overall picture of your community and the nation – NOT to find out about you personally.)
> AT APPROPRIATE MOMENTS, ENCOURAGE RESPONDENT TO GET RECORDS.
> PRESS "ENTER" TO CONTINUE

The new text was adapted almost word-for-word from text developed by Moore and Loomis (2002), who also sought a way to reduce income amount item nonresponse. Their strategy was to invoke "the sorts of factors which previous research (e.g., Gerber, Crowley, and Trencher, 1999) have found to be important 'legitimizers' of requests for sensitive information" [p4]. Moore and Loomis experimentally tested the impact of the addition of the introduction and found that it significantly reduced nonresponse to income amount items. It appears to have done so here as well, although the evidence is admittedly indirect and circumstantial – it makes intuitive sense, it has worked in other contexts, there are no other obvious explanations, and the improvement in SIPP 2004 is of the "right" form: reduced refusals. No similar effect is apparent in the wave 1 item nonresponse results for asset amounts (Moore, 2006), a fact that is somewhat inconvenient for the present conclusion. But the asset amounts section is confounded with other new features

that might have masked any positive trends – in particular, nonresponse follow-up procedures, with their potential to affect interviewers' behavior – thus making it a less pure section in which to test for any effects of the new introduction.

### 4.4. Research Limitations

It is rare that any research offers unassailably definitive findings, and the current project is certainly no exception. Several factors limit the confidence with which conclusions may be drawn from this research. Chief among them, perhaps, is the fact that the investigation had to use a "natural experiment," as opposed to a designed one. As a result, the analysis must rely on an underlying assumption of equivalent survey conditions, and the absence of confounding factors affecting the comparisons of interest, when in fact the survey conditions were not controlled. In fact, many differences, in addition to the questionnaire differences of interest here, are easily identifiable:

(a) the great many other questionnaire design differences which were confounded with those of interest to this investigation;
(b) the use of different interviewer pools (overlapping, to be sure, but by no means identical);
(c) different sample designs;
(d) different immediate survey contexts – e.g., the 2001 advance letter advised respondents that "answering the questions is voluntary"; the 2004 letter tried to reduce the likelihood of complete nonresponse by presenting voluntary cooperation in a different light ("you may decline to answer any particular question") which may have affected item nonresponse proclivities; and
(e) different socio-cultural milieus (e.g., all of the 2001 wave 1 interviews, and almost all of the 2001 wave 2 interviews, took place before the 9/11 terrorist attacks).

No doubt there are many other differences as well, all of which are ignored in this investigation, and all of which could affect the 2001-2004 comparisons in unknown ways. Another important limitation is the use of the unedited and unweighted "TransCASES" data files; edited and weighted data may yield different results. While it is necessary to note these caveats about potential confounds, it is just as essential to note that there is no indication that any of them actually *did* affect the results in important ways, or the conclusions drawn from them.

Finally, it is useful to acknowledge once again that the analysis of the benefits of the use of dependent questions to follow-up an initial nonresponse assumes an equivalence between the response to the follow-up procedure and an initial dollar amount report. On the one hand, it seems quite reasonable to treat a respondent's confirmation of a prior wave amount (or his/her report of a corrected value) as having negated his/her initial nonresponse. But whether that process actually produces income data of equivalent quality is not quite so clear.

## 5. Conclusions and Discussion

The SIPP 2004 wave 2 questionnaire implemented new follow-up procedures in an attempt to reduce nonresponse to the survey's "general income" amount items. The results presented in this paper suggest that those procedures met with considerable success – they were very effective at obtaining information from those who initially failed to provide an amount report. As a result, compared to the 2001 panel, the final rates of nonresponse for those items were significantly reduced in the 2004 panel, in most cases to less than half the rate that had been experienced in the past. In this sense, certainly, the new procedures were an unequivocal success.

However, as noted before, it is important to keep in mind that the benefits of the new dependent follow-up procedures are not without costs. Particularly troubling are the very large increases in initial nonresponse in wave 2, which are almost certainly due to those procedures. This appears to be a matter of bringing interviewer behavior under better control, so that they allow the natural response process to unfold before turning to the follow-up options. Whether this is possible, given the understandable natural pressures on interviewers to access those procedures with little delay, and given the apparent need to "unlearn" what is by now perhaps an habitual practice (encouraged, unfortunately, by some off-target early training), remains to be seen. Other costs of these procedures are more practical and operational in nature, though no less important – dependent interviewing procedures add complexity to a survey instrument, and complicate data analysis, and require that much additional effort be devoted to such activities as the extraction of the appropriate data from one survey wave, and the transfer of those data to the next wave.

The significant reduction in wave 1 item nonresponse was an unexpected finding of this research. Perhaps it should not have been unexpected. Eliminating several "usual suspects" leaves us with the suggestion that the revised introduction had the effect which, after all, was intended – namely, that it reduced somewhat respondents' reluctance to reveal their income amounts. As noted, there are many reasons why caution about this conclusion is appropriate. But it does at least offer the hope that positive change in item nonresponse outcomes does not necessarily require difficult or complex modifications to existing procedures, but may instead only require changing a few words.

This research highlights several obvious candidates for future research, some of which have already been alluded to. Are there testable hypotheses regarding factors other than the new introductory text which could explain the reduction in wave 1 nonresponse? Can interviewer behavior be modified? What is the impact, if any, of the wave 6 refresher training on their tendency to be too quick to access dependent information? What impact does the ability to access prior wave amount data have on other aspects of data quality, such as measurement error? In concert with the present study, which does suggest some clear positive benefits to the new procedures, additional research on these and other topics will allow a more thorough cost-benefit evaluation.

**Acknowledgments**
A great many people provided assistance with various aspects of the work presented in this paper. At the risk of inadvertently omitting key names, I specifically note the following: Anna Chan, Julia Klein Griffiths, and Joanne Pascale, my immediate colleagues for many years on the SIPP Methods Panel project; my Demographic Surveys Division (DSD) colleagues on that same project, especially the late, great Pat Doyle and her staff – Nancy Bates, Tim Gilbert, Elaine Hock, Heather Holbert, and Aniekan Okon; and reviewers who provided many useful comments on earlier drafts of this paper – in addition to several people mentioned above, that list includes Judy Eargle (DSD), John Hisnanick (Housing and Household Economic Statistics Division), and Somonica Green (Field Division).

**References**

Bruun, M. (2005), "Fifth Wave SIPP Observation," unpublished U.S. Census Bureau memorandum to Jeff Moore, Statistical Research Division, July 12, 2005.

Davis, J. (2005), "Observation Notes of Personal Interviews for the SIPP Wave 5: Evanston, Skokie, and Wilmette, IL," unpublished U.S. Census Bureau memorandum to Doug Geverdt and Kurt Bauman, Population Division, June 7, 2005.

Doyle, P., Martin, E., and Moore, J. (2000), "Methods Panel to Improve Income Measurement in the Survey of Income and Program Participation," Proceedings of the Section on Survey Research Methods, Alexandria VA: American Statistical Association, 953-958.

Gerber, E., Crowley, M., and Trencher, S. (1999), "Identity Thieves, Warrantee Cards and Government Surveys: The Ethnography of Personal Information Management," paper presented at the International Conference on Survey Nonresponse, Portland, OR, October 28-31, 1999.

Gilbert, T. (2005), "SIPP 2004 Panel Wave 4 Observation," unpublished U.S. Census Bureau report, May 6, 2005.

Moore, J. (2004), "SIPP 2004 Panel Wave 2 Observation Report," unpublished U.S. Census Bureau report, September 27, 2004.

Moore, J. (2006), "The Effects of Questionnaire Design Changes on Asset Income Amount Nonresponse in Waves 1 and 2 of the 2004 SIPP Panel," Washington, DC: U.S. Census Bureau, Research Report Series (Survey Methodology #2006-01), issued January 4, 2006.

Moore, J. and Griffiths, J. (2003), "Asset Ownership, Program Participation, and Asset and Program Income: Improving Reporting in the Survey of Income and Program Participation," Proceedings of the American Statistical Association (cd-rom), Alexandria VA: American Statistical Association.

Moore, J. and Loomis, L. (2002), "Reducing Income Nonresponse in a Topic-Based Interview," Washington, DC: U.S. Census Bureau, Research Report Series (Survey Methodology #2002-06), issued March 11, 2002.

Moore, J., J. Pascale, P. Doyle, A. Chan, and J. Klein Griffiths (2004). "Using Field Experiments to Improve Instrument Design: The SIPP Methods Panel Project." In S. Presser, J. Rothgeb, M. Couper, J. Lessler, E. Martin, J. Martin, and E. Singer (eds.), Methods for Testing and Evaluating Survey Questionnaires. New York: Wiley, Chapter 10, pp. 189-207.

Moore, J., Stinson, L., and Welniak, E. (2000), "Income Measurement Error in Surveys: A Review," Journal of Official Statistics, 16: 331-361.

U.S. Census Bureau (2001), "SIPP Users' Guide," Washington, DC: U.S. Census Bureau.

U.S. Census Bureau (2004), "SIPP 2004 Panel Wave 2+ Guide for Training Field Representatives," Washington, DC: U.S. Census Bureau, May 2004.

## Tables

### Table 1: Nonresponse to Monthly Income Amount Items for Selected "General Income" Programs in Wave 1 and Wave 2 of the 2001 and 2004 SIPP Panels

[Data source: TransCASES data files, unweighted]

| General Income Program | Income Amount Item Nonresponse Rate (number of amount questions asked) | | | | |
|---|---|---|---|---|---|
| | Wave 1 (combined across m1-m4) | | Wave 2 (combined across m1-m4) | | |
| | 2001 | 2004 | 2001 | 2004 | |
| | (no attempt to follow-up nonresponses) | | | Initial n-r (pre d.i.) | **Final n-r** (post d.i.) |
| (01) Social Security | 20.0 (48,643) | 15.8 (62,087) | **23.4** (39,522) | 56.0 (59,787) | **11.4** |
| (03) Federal SSI | 9.8 (7,227) | 8.8 (6,456) | **14.5** (5,916) | 40.8 (6,474) | **7.8** |
| (05) State Unemployment | 7.0[a] (2,123) | 8.0[a, b] (4,074) | **11.8** (1,527) | 24.1 (2,676) | **7.7**[b] |
| (08) Veterans' Compensation/Pensions | 13.5 (2,915) | 11.6 (4,380) | **22.0** (2,406) | 55.5 (4,292) | **9.5** |
| (10) Workers' Compensation | 16.0 (982) | 13.0 (1,130) | **22.3** (620) | 37.7 (1,046) | **7.9** |
| (27) Food Stamps | 5.3 (8,841) | 4.3 (13,768) | **6.9** (7,249) | 28.8 (14,243) | **2.8** |
| (28) Child Support | 8.8 (6,264) | 6.5 (9,033) | **10.9** (5,305) | 38.0 (8,746) | **4.7** |
| (29) Alimony | 14.7[a] (700) | 16.6[a] (795) | **23.7** (590) | 44.3 (761) | **8.4** |
| (30) Company/Union Pension | 21.0 (13,894) | 17.2 (18,275) | **27.9** (10,941) | 58.4 (18,326) | **14.8** |
| (31) Federal Civil Service Pension | 21.5 (2,165) | 16.2 (3,048) | **28.0** (1,597) | 61.6 (3,041) | **11.0** |

Notes:

(1) Nonresponse estimates in essence show the proportion of all amount questions, accumulated across all respondents, which failed to elicit an amount report (see text).

(2) Matching superscripts indicate estimates that do not differ statistically; all other differences (across panels within a wave, and within a panel across waves) are statistically significant at the $p<.10$ level or beyond.

(3) "pre d.i." = initial rate of nonresponse, before implementing the dependent interviewing follow-up procedures; "post d.i." = rate of nonresponse remaining after implementing the dependent interviewing follow-up procedures.

## Table 2: Percent Distribution of Initial Responses to SIPP 2004 Wave 2 "General Income" Amount Items, by Rotation Group

[Data source: TransCASES data files, unweighted]

Part a: All GI Sources Whose Amounts are Captured in MNTHAMT15, Combined [see note]

| Initial Response | Rotation Group | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| Reported a $ amount (inc. 0 or "None") | 60.3 | 49.7 | 43.5 | 40.7 |
| **"L" ("same as last time")** | **24.7** | **39.0** | **45.2** | **49.1** |
| Don't know / Refused | 15.0 | 11.3 | 11.3 | 10.2 |
| Total | 100% | 100% | 100% | 100% |
| (n) | (7,662) | (8,028) | (7,929) | (7,701) |

Part b: Food Stamps

| | | | | |
|---|---|---|---|---|
| Reported a $ amount (inc. 0 or "None") | 80.2 | 73.7 | 72.9 | 63.4 |
| **"L" ("same as last time")** | **14.3** | **22.2** | **23.6** | **34.1** |
| Don't know / Refused | 5.5 | 4.1 | 3.5 | 2.5 |
| Total | 100% | 100% | 100% | 100% |
| (n) | (935) | (943) | (1,004) | (1,030) |

Part c: Child Support

| | | | | |
|---|---|---|---|---|
| Reported a $ amount (inc. 0 or "None") | 75.9 | 64.5 | 61.4 | 58.0 |
| **"L" ("same as last time")** | **18.7** | **29.2** | **35.0** | **38.0** |
| Don't know / Refused | 5.4 | 6.3 | 3.6 | 4.0 |
| Total | 100% | 100% | 100% | 100% |
| (n) | (555) | (631) | (634) | (645) |

Notes:
(1) The MNTHAMT15 screens capture income amounts from 39 of the approximately 50 general income sources, including the following: Social Security, Railroad Retirement, Federal SSI, state SSI, veterans' compensation and pensions, Black Lung payments, employer disability payments, Energy Assistance, General Assistance/General Relief, foster child care payments, alimony, a pension from a company or union, Federal Civil Service pension, US military retirement, National Guard/Reserve retirement, state government pension, local government pension, income from a paid up life insurance policy or annuity, estates/trusts, GI Bill benefits, food assistance (other than Food Stamps), clothing assistance, short-term cash assistance, and a handful of other rare and/or miscellaneous categories. Only a very few major transfer programs – e.g., TANF benefits, Food Stamps, and child support – are <u>not</u> captured in MNTHAMT15.