

**THE SURVEY OF INCOME AND  
PROGRAM PARTICIPATION**

**TESTING ALTERNATIVE HOUSEHOLD  
ROSTER QUESTIONS FOR THE SURVEY  
OF INCOME AND PROGRAM  
PARTICIPATION**

**No. 164**

**D. Cantor and C. Edwards**

**Testing Alternative Household Roster Questions  
For The Survey Of Income And Program Participation**

Final Report

David Cantor  
Colleen Edwards

Submitted to the Bureau of the Census  
under BLS Contract #J-9-J-8-0083

May 5, 1992

## TABLE OF CONTENTS

<u>Chapter</u>		<u>Page</u>
	EXECUTIVE SUMMARY .....	i
1	RESEARCH ISSUES .....	1
2	RESEARCH DESIGN AND RESPONSE RATES.....	3
	2.1 Sample Design.....	4
	2.2 Recruitment, Training and Interviewing Procedures .....	5
	2.3 Response Rates .....	6
3	RESULTS .....	7
	3.1 Comparison of Males in Different Demographic Groups .....	7
	3.2 Qualitative Description of Non-Usual Residents .....	10
4	CONCLUSIONS AND RECOMMENDATIONS .....	13
	4.1 Incorporating Residency into Respondent Definitions .....	13
	4.2 Future Research .....	14
	REFERENCES .....	17

## LIST OF EXHIBITS AND TABLES

		<u>Page</u>
Exhibit 1	Household Enumeration Questions Used For the Survey of Income and Program Participation.....	18
Exhibit 2	Household Enumeration Questions Based on Overnight Residence at Unit (Version 2) .....	19
Table 1	Demographic Distribution of Residents Listed on the Household Roster by Questionnaire Version.....	20
Table 2	Demographic Distribution of Residents Listed on the Household Roster by Questionnaire Version .....	21
Table 3	Demographic and Residency Information for Persons Reported as Not Usually Living in the Unit by Questionnaire Version and Household .....	22

## Executive Summary

The purpose of this project was to test whether an alternative procedure to roster the household would reduce undercoverage on the Survey of Income and Program Participation (SIPP). The current procedure, which is identical to that used on most other household surveys, administers questions to a household respondent which require him/her to list members of the unit who "live or stay there". This approach has been criticized because terms such as "live" or "stay here" are ambiguous when applied to individuals who are either not legally attached to any particular household member or who actually reside in more than one unit.

The alternative roster questions tested on this project asked respondents to report all individuals who had stayed in the unit, for at least one night, since July 1. The hypothesis was that this procedure would increase the number of persons listed on the household roster who have been found to be undercounted on the SIPP (primarily nonwhite males). This is based on the assumption that, when compared to the current SIPP procedure, the alternative approach will increase the number of persons who have an ambiguous attachment to the unit. If this is true, then at least part of the undercoverage problem could be eliminated by adopting a procedure based on direct residency rules.

To test this hypothesis, two 4-block areas in southeast Washington D.C. were selected for study. These areas were selected on the expectation that the households would have a disproportionate number of individuals who have a high chance of not being enumerated with the current SIPP procedures. The housing units were listed in each area and a sample of 53 pairs of neighboring units were selected. Attempts were made to administer the SIPP procedure (Version 1) to one unit of each pair and to administer the alternative procedure (Version 2) to the other half of the pair. Once eliminating ineligible units, the study achieved an overall response rate of 70%. A total of 27 complete pairs were interviewed. The sample was predominantly black.

The results provide support for the argument that the alternative approach will result in better coverage. Overall, Version 2 resulted in twice as many males age 18-60 as Version 1. The difference, however, could not be directly attributed to differences in the two procedures. This leaves open the question whether the difference appeared simply by chance. In addition, however, debriefing information from both versions indicated that a number of persons were identified as "not usually living" at the unit, but who clearly had ties to a member of the household. These individuals, primarily young males, would not have been listed on the roster when using the normal SIPP procedure.

The report concludes by recommending that further research should be devoted to developing a procedure to roster the household that is based on the residence criteria. The advantage of this approach is that it reduces the role of respondent judgement when deciding on who should be listed on the initial household roster. Decisions on eligibility for the survey, therefore, will shift from the respondent to the survey designer/analyst. The disadvantages of this approach are that it may increase other types of response error, increase response burden and complicate estimation of population parameters. Specific recommendations for future research include: 1) investigating the alternative procedure with a larger sample of individuals, 2) more detailed examination of respondent interpretation of the current SIPP roster questions, 3) investigating related patterns of undercoverage with Decennial Census data, and 4) more specific examination of the estimation issues that would accompany the adoption of the alternative procedure.

## 1. Research Issues

As part of the research on undercoverage for the Survey on Income and Program Participation (SIPP), a small survey was conducted in two areas of Washington D.C. The purpose of the survey was to assess whether changing the initial household roster questions on the control card would provide significant increases in the number of males (especially young black males) who are listed on the household roster. The purpose of this report is to describe the results of this survey and make recommendations for future research.

The issue of undercoverage on the Decennial Census and most household surveys is becoming an increasingly important problem. For a number of surveys, this problem results in missing many of the most important portions of the target population. This is especially the case for the SIPP, which is specifically targeting individuals who participate on government assistance programs. Information on welfare reciprocity, food stamps or other "safety net" programs requires contacting those individuals who are least likely to be listed on a household roster for most household surveys.

Hainer, et. al., (1988) summarize extant research on this problem by classifying the error into two different categories. First, there is the tendency of certain types of respondents to be wary to list household members because it may jeopardize benefits. A large proportion of government program recipients qualify only if the household income is below a certain level. These recipients, therefore, are less likely to report the residence of an individual who has not been reported to program administrators. Similarly, lower income individuals also have a higher likelihood of being late on bills. This increases the suspicion on the part of household members when strangers approach them for information on household residents because it may be some type of bill collector or law enforcement officer.

A second type of error is directly related to the types of questions that are asked on the household roster. These items (Exhibit 1) include an initial question on the

"...names of all the persons living or staying here...".

These items are then followed by a series of questions that ask if the respondent had missed a number of other types of potential household members, such as small babies, lodgers, boarders, etc.. and whether each person listed "usually" lives there.

As summarized by Hainer, et. al., these questions are ambiguous. These items do not provide clear guidance on who should be reported because key terms, such as "staying" and "usually living here" are not explicit for a number of different types of living situations. Examples of individuals where this ambiguity would likely arise include:

- members of a couple who are not living together, but spend large amounts of time in one apartment,
- sons/daughters of divorced couples with shared visitation,
- highly mobile individuals who spend time in several residential locations on a regular basis,
- individuals who are away because of work for long periods of time and are irregularly resident anywhere,

When applying the concept of "household" and "usually lives here" to these individuals, there is quite a bit of room for interpretation.

The problem of interpretation of the roster questions is not limited to "undercounted" household residents. Researchers have argued, however, that the "household" situation for underenumerated individuals should lead to a higher frequency of problems which result from these ambiguities. Most U.S. households fit into the stereotype definition that is implied by the household roster questions (i.e., intact families where household members do not stay in other household units). The error resulting from these ambiguities for these individuals, therefore, is less when compared to individuals where family groups and residency patterns are not defined in these terms (e.g., see Table 4 in Hainer, et. al., 1988).

As documented in a number of sociological and anthropological studies (e.g., Wilson, 1987), there has been an increasing trend in unusual living situations among a small, but highly significant, segment of the household population. This trend corresponds fairly closely with rises in undercoverage in the Decennial Census.<sup>1</sup> If the Decennial Census is having this problem, it can only be assumed that household surveys, like SIPP, are experiencing coverage problems at least as great (Shapiro and Kostanich, 1988). One would expect, therefore, that if there are no changes in the way households are defined by social researchers, the problem of coverage will not only be significant, but will likely increase in the future.

---

<sup>1</sup> This increase in undercoverage is occurring despite an increase in the effort by the Bureau to reduce this problem.

The purpose of this research was to address this second type of error. This was accomplished by comparing the current SIPP household roster questions to items based on more sharply defined criteria. As will be described below, these criteria were based primarily on residential patterns over a well defined reference period. Examples of these items include: 1) who stayed in the house overnight since July 1, 2) how many nights the person stayed, and 3) why the person was no longer staying in the unit. This alternative approach is based on the assumption that a significant part of the undercoverage on the SIPP (and other household surveys) is due to problems with respondent interpretation of phrases such as "live or staying here" or "usually live here". In particular, the tendency is for respondents to apply an overly restrictive definition of who should be listed on the household roster. This, in turn, results in significant undercoverage of individuals who have ambiguous ties to the household unit.

The alternative procedure tested here sharpens the meaning of the questions, while also broadening the scope. This leaves less room for respondent judgement when deciding on who should or should not be listed on the roster. One would expect, therefore, that if undercoverage problems are due to the errors described above, this change in the procedure to roster the household should significantly improve coverage of the population.

The alternative procedure also shifts much of the burden of determining eligibility for the interview from the respondent to the survey designer/analyst. This requires that the survey analyst and designer decide on explicit characteristics that should be used to define eligibility criteria. In principle, this should be a better practice than leaving these rules up to the respondent. In the remainder of this report, the design, results and conclusions from this research are described in more detail.

## **2. Research Design and Response Rates**

The design for the study consisted of administering two different versions of the household roster questions to pairs of neighboring household units. The units were selected from two small areas in Washington D.C. that were suspected to have high rates of undercoverage. One set of roster questions were identical to that used on the SIPP (hereafter referred to as Version 1). This procedure was briefly described in Section 1 above (see Exhibit 1). The second procedure (hereafter referred to as Version 2) was designed to elicit household roster information for all individuals who had stayed in the unit at least one night since July 1 (Exhibit 2).

The Version 1 questionnaire (Attachment 1) consisted of: 1) the SIPP household roster questions, 2) demographic questions taken from the SIPP control card, 3) a short set of employment questions, and 4) debriefing questions. The debriefing questions asked the respondent to list any other persons, not listed on the original household roster, that had stayed at the house for at least one night, since July 1. In addition, the debriefing items included a number of questions on residency patterns since July 1, for all persons listed on both the original and debriefing household roster. Both sets of debriefing questions were used to measure the extent to which the original Version 1 roster may have changed, if the rules were based on simple residency.

The Version 2 questionnaire (Attachment 2) consisted of: 1) the experimental household roster question, 2) items on residency patterns since July 1, 3) demographic questions from the SIPP control card, 4) a short set of employment questions identical to those on Version 1 and 5) debriefing questions. The debriefing questions were designed to have the interviewer review each person listed on the household roster and ask the respondent to identify which individuals "usually" live at the unit. This was then followed up with a question on the criteria respondents used to decide on usual residence.

These questionnaires were used to assess the utility of the experimental procedure in two different ways. First, a comparison of the sex ratios across Version 1 and Version 2 households was made. If Version 2 produced more males listed as "non-usual" residents then this would provide support for the new procedure. Second, the type of individuals who were listed as "non-usual" residents was examined with respect to residential patterns. If asking about residence since July 1 produced individuals who have ambiguous ties to the household, then this would also provide evidence for the utility of the experimental procedure.

## **2.1 Sample Design**

The sample was selected from two four-block clusters in southeast Washington D.C.. The particular areas were selected on the advice from a consultant familiar with Washington D.C. and the problems of enumerating households in a survey context. All households in the two clusters were listed by Census enumerators. Once listed, neighboring housing units were formed into pairs of addresses. A sample of 65 pairs were selected for the study.

## **2.2 Recruitment, Training and Interviewing Procedures**

Interviewers were recruited from residents of the DC area who were willing to work in the sampled areas. Three interviewers were hired, primarily based on references and an in-person interview with the field supervisor. A two day training was conducted with the interviewers. This training included a general review on interviewing techniques, household contact procedures and review of questionnaire specific items.

Two Census Bureau interviewers were also trained on the procedures and were used to complete refusal conversion. This training was conducted in one day, since only study specific procedures had to be reviewed.

Interviewers were trained on both versions of the questionnaire and were randomly assigned households for both versions. The interviewing procedures included:

1. sending a letter to each housing unit introducing the residents to the study and the sponsor,
2. an interviewer making contact with a member of the household. The interviewer asked to speak to the person who owned or rented the home. If that person(s) was not there, any person who was at least 18 years old was eligible to answer the questionnaire,
3. the interviewer explaining the conditions of the study. This included:
  - participation was voluntary
  - all information was confidential, and

If the respondent agreed to these conditions, he/she would sign a consent form

4. the interviewer requesting to tape record the interview,
5. administering the interview.

Three problems were encountered during the course of the study. First, access to the units was extremely difficult because most were located in a multi-unit structure which had a locked

entrance ( and no resident manager). Interviewers were reluctant to gain access to units by calling neighboring apartments because of the dangerousness of the area.<sup>2</sup>

Second, once gaining access and contacting the unit resident, interviewers were not successful in gaining permission to tape record interviews. We suspect this was due, in part, to the suspicious nature of the respondent population. It was also due, in part, to the inexperience of the interviewers. None of the Westat interviewers had previous interviewing experience. We strongly suspect that the interviewer was reluctant to push the tape recording once gaining access to many of the apartment buildings and talking respondents into co-operating with the study. More experienced interviewers may have been able to make a more convincing presentation to gain consent.

Third, one of the interviewers was found to have falsified a large number of interviews. Once discovering this, a new interviewer was trained and assigned all those cases that were completed by the original interviewer. If the case could not be verified or a complete could not be obtained by the new interviewer, it was treated as missing data.

### **2.3 Response Rates**

The interviews were completed over a two month period between August 20 and October 20. Out of the 130 addresses that were listed:

- 8 were defined as out-of-scope because an interviewer reported being robbed in the building,
- 14 units were found to be vacant,
- 2 were for addresses that did not exist,

This left 106 eligible addresses where one could reasonably expect to complete an interview.

Out of this 106, a total of 77 interviews were obtained. The remaining non-response was due to: 1) six refusals, 2) 14 cases no-one in the household could ever be contacted, 3) three cases where the wrong survey was administered, 4) five cases where the interview could not be verified by the field supervisor and 5) one case that was lost in the mail.

---

<sup>2</sup> There were occasional shootings reported in several of the areas immediately adjacent to sampled areas.

The missing data was disproportionately located within Version 1 (62% response rate for Version 1, 81% response rate for Version 2). One important reason for this was that all of the instances where the wrong version was administered were in Version 1 households. This version also had slightly more cases where the unit was found vacant, there was no contact made and the respondent refused to co-operate.

### **3. Results**

In this section the results of the survey are presented. This discussion evaluates the experimental procedure in two different ways. In the first section, a quantitative comparison is made between the number of males and the number of females in different demographic categories. The hypothesis is that if the new procedure increases coverage, then the household roster should contain proportionally more adult males. In the second section a qualitative evaluation is made of the types of persons who were listed as "not usually living" in the unit but who stayed for at least one night since July 1. One would consider the experimental procedure a success if it captured individuals who may not usually live in the unit, but did spend a significant amount of time there since July 1.

Table 1 provides the demographic characteristics, by version, for all interviewed households (n=77 households). It can be seen that the individuals listed on the household rosters were predominantly black. About a third were married and approximately 25% had attended some college. There are differences in age-sex distributions across the two questionnaire versions. These differences, however, should be interpreted with caution, since there was an imbalance between the number of completes across the two questionnaire versions.

The analysis described in Section 3.1 below is based on only those cases where both households of the sampled pair completed the correct questionnaires (n=54 households). It is with this reduced base that comparisons between questionnaire versions should be made. The analysis described in Section 3.2 is based on all 77 of the cases interviewed for the study.

#### **3.1 Comparison of Males in Different Demographic Groups**

Table 2 provides selected statistics on the basic demographic and residency information for only those cases where both units of the original pair completed an interview (N=27 for each questionnaire version). These data exclude the information collected during either the Version 1 or

Version 2 debriefings. For Version 1, there were a total of 65 people listed on the household roster (i.e., people who usually live there). This compares to 72 individuals who were listed on the experimental version (Version 2) who had stayed at least one night since July 1. As one would expect given the sampled areas, the vast majority of the respondents were black (94% for Version 1 and 100% for Version 2).

The sex ratios across the two questionnaire versions are different and in the expected direction. For Version 1, the proportion of males is 40% , while for Version 2 this proportion is 46%. The primary reason for this difference is that the proportion of middle age males (i.e., age 18-60) who were enumerated in the Version 2 questionnaire is twice as large as that enumerated in Version 1 (31% vs. 15%).<sup>3</sup> In terms of absolute numbers, this difference translates into 10 males age 18-60 for Version 1 and 23 for Version 2.

Why are there more middle aged males in Version 2? The hypothesis is that by asking about overnight residency, rather than about who lives in the unit, respondents would include individuals that may have ambiguous ties to the unit. The observed pattern in the sex ratios for middle age residents is consistent with this hypothesis. Closer inspection of the data, however, does not directly support this conclusion. This is evident when disaggregating the debriefing data by sex and age. Of those 72 individuals enumerated on the Version 2 questionnaire, 14 were classified in the debriefing as not "usually" living there. Slightly more than half of these 14 were females. Once eliminating these 14 from the Version 2 data, therefore, the proportion of males age 18-60 is still about twice as large as for Version 1. In other words, once eliminating those persons who seemed to have been added to the Version 2 roster because of the experimental procedure, the proportion of males age 18-60 remains unchanged.

A similar pattern is evident from the debriefing data for Version 1. In this debriefing, respondent's were asked to list all individuals who had stayed in the unit at least one night since July 1. For each person listed, follow-up questions on residency patterns, identical to what were collected on Version 2, were administered. A total of 10 extra people were listed on this part of the Version 1 debriefing. The proportion of males age 18-60 for these individuals was identical to those enumerated as "usually" living in the household.

---

<sup>3</sup> If one computes a two tailed t-test, assuming simple randomsampling and using the total number of persons enumerated as a base (i.e. 65 for Version 1 and 72 for Version 2), this difference is significant at the .05 level.

The differences between Version 1 and Version 2 with regard to males 18-60 years old, therefore, is not clear. It could, of course, be due to chance. That is, despite our best efforts to match households on an area basis, Version 2 was administered to households with a larger proportion of males in the 18-60 age group. It may, however, also be due to the order in which the residency questions were asked. The way respondent's understood the questions in the two different versions could have been affected by the order in which the questions were presented in the two different questionnaire versions. For Version 1, respondents interpretation may have been narrowed by the fact that the question on residents since July 1 was asked after the "usually live here" question. This may have occurred because:

- respondents thought that since he/she had already been asked about people who usually live there, the question on residents since July 1 should be restricted to individuals who clearly are not "usual" residents. That is, individuals who stay at the unit on a regular, but very sporadic basis, simply did not come to mind. Only those who were clearly there for a brief, one-time only, occasion were retrieved from memory,
- the initial questions may have sensitized the respondent to revealing too much about residents of the unit. This may have restricted the types of individuals the respondent was willing to report later in the interview,
- the amount of effort to recall all individuals who stayed there was not exerted. Since the residence questions were at the end of the control version, the respondent may have simply aborted searching memory because of fatigue or simply a desire (conscious or unconscious) to terminate the interview.

In any of these cases, the implication is that an approach that based the roster on simple residence rules should produce more individuals that are typically undercounted.

This explanation, based on the order/context of the two questionnaires, was explored by examining the residential patterns of those individuals who were reported as "usually" living at the unit by questionnaire version. If the residents listed for Version 2 were not as "attached" to the household as in Version 1, then this would provide some evidence that the initial Version 2 household roster provided a "different" type of resident. Residency patterns were examined using the questions which asked about: 1) the number nights the person had slept overnight at the unit and 2) the number of different residences the person stayed in during the reference period.

Inspection of these data do not reveal any differences by questionnaire version. For those who were identified as "usually" living in the unit, there was not a significant difference in either of these variables across the two different questionnaire versions. Almost every person, for both versions, claimed to have stayed in the unit for the entire reference period and in only one unit. If the

above explanation were true, one would have expected that Version 2 residents would have slept fewer nights at the unit and would have stayed in more residents during the reference period.

In summary, the relevant quantitative comparisons between the two questionnaire versions produce results that are consistent with the hypothesis that the experimental version increases coverage of males age 18-60. The explanation for this result, however, cannot be directly linked to the experimental procedure. This leaves open the question of whether these results are due to chance or some other explanation related to the field procedures.

### 3.2 Qualitative Description of Non-Usual Residents

The quantitative data presented in section 3.1 is limited by the extremely small sample sizes. A second way to evaluate the experimental procedure is to directly examine those individuals listed as staying overnight but not "usually" living in the household. As described above, a number of demographic and residential questions were administered for each person listed on each household roster. This included: 1) relationship to head of household, 2) how many nights the person stayed since July 1, 3) whether the person expected to stay over the next 4 months, 4) the reason why the person left and 5) the criteria used to define whether the respondent "usually" lived in the house (Version 2 only). Examination of these data for persons who were listed as staying overnight but "not usually living" in the household provides insight on the utility of experimental procedure. If the only type of resident listed are those who are staying for extremely short periods of time and have no clear attachment to the household, then one would not conclude that the new method produces significant increases in coverage. Alternatively, if "hard to enumerate" individuals are identified, then the alternative procedure would be worth considering.

The results, although based on a small sample size, provide evidence that the alternative procedure did result in listing several residents who are "hard to enumerate". Table 3 provides these data, by household membership and questionnaire version. The data listed here are taken from all the cases collected from each completed household unit. These are not restricted to only those cases where both cases of the original pair were completed.<sup>4</sup> As can be seen, for Version 2, over half of the 15 non-usual residents are in a single household. The respondent for this household reported that these individuals were simply visiting for a very short period of time. Similarly, for household 1, relatives

---

<sup>4</sup> If one restricts the universe to only those cases where there were matching pairs of completed interviews, one case from Table 3 drops out. This is the case included as Household #6 for Version #2.

of the respondent were reported as staying for a short period of time for the primary purpose of visiting. Criteria used to define "usual residence" in both cases was whether the person had his/her own place to live.

In households 3, 4, 5 and 6, however, the situation is more ambiguous. In household 3, the separated husband of the respondent had stayed for 5 days and was still staying at the residence at the time of the interview. The respondent reported that although the person would not stay for an additional 4 months, the date of departure was not known. This seems to be a situation, therefore, where the husband is not defined as living in the unit, but likely spends a significant amount of time there. In this case, the wife defined the husband as not usually living there because he had his own place.

In Household 4, the brother of the respondent was not defined as "usually living" at the unit, despite the fact that he had stayed there for 30 days. The brother was reported as having just moved out of an apartment and was in search of another place to live. The respondent did not know exactly when the brother was leaving.<sup>5</sup> As with the other households, this person was defined as not usually living at the unit because of the expectation that he would go back to his own apartment (once he found one).

In Household 5, a grandson of the respondent had been staying in the unit for 8 days and was still staying there at the time of the interview. The respondent reported that it was also expected that he would be staying for the next 4 months. When asked why the person was not considered as "usually" living there, the respondent said that he was expected to return to his father's house at some point in the future.<sup>6</sup>

Household 6 provides a slightly less ambiguous case. This is an 8 year old girl who stayed with her grandparents for a 30 day period. The girl was no longer staying with the grandparents at the time of the interview. From the information on the questionnaire, this seemed to be a visit that was clearly defined and had a logical start and end point. On it's face, this type of situation should not lead to enumeration problems, since the girl seemed to be clearly attached to her parent's house. In reality, however, it is not clear whether she would have been picked up in either the grandparent's or parent's home when using the usual SIPP roster procedures.

---

<sup>5</sup> She reported expecting the move sometime in the next month, but could not give a specific date.

<sup>6</sup> If the person was staying at the house 4 months beyond the interview, the respondent was not asked about a specific date when the person would be leaving.

A total of 10 individuals in 6 households were listed in the Version 1 debriefing as having stayed overnight in the unit who were not listed on the original household roster as "living there". Of these, 2 individuals could be identified as posing potential enumeration problems. One of these persons was reported as having stayed there for a significant period of time (2 months) but was no longer living there. This was a male, 17 years old, who visited the respondent's son every summer. It is understandable why the respondent did not list this on the original household roster, since he was not living there at the time of the interview. Nonetheless, the significant time he spent, on a regular basis, at two different households may have posed significant enumeration problems if the interview had taken place earlier.

The second example was a household where the respondent reported a male friend who was currently living at the unit and expected to stay at the unit for the next 4 months. This individual had one other household where he had stayed during the reference period.

It is interesting to note that in most of the ambiguous cases identified, the resident was male and between the ages of 17 and 36. This reinforces the notion that these individuals are part of the group that is typically undercounted. It also might explain the results discussed in Section 3.1. It may be the case that some portion of the differences in the sex ratio between the questionnaire versions is due to the effect the Version 2 procedure has on the number of young males listed on the roster. The differences at the older age groups may have been due to some other reason, including sampling variation or unanticipated differences due to the procedure.

It should also be noted that in most cases, the "ambiguous" residents were currently staying at the unit at the time of the interview. Very few individuals who were not currently there and who spent significant periods of time at the unit were reported. This could be due to a number of reasons, including the reluctance of respondents to report these individuals or memory failure.

In summary, these data indicate that asking the alternative household roster questions does provide information on individuals who are typically undercounted. Those individuals identified were primarily those who were staying at the unit at the time the interview. The procedure also produced a large number of persons who were simply temporary visitors. The implications of these results for future research are discussed in more detail below.

#### **4. Conclusions and Recommendations**

The results described above indicate that asking respondents about location over a well defined reference period has clear potential for reducing undercoverage. The quantitative results described above are consistent with this conclusion. Closer inspection of the data, however, cannot directly link the procedure to these quantitative patterns. Given the sample size for this project, there is some chance that the difference was simply due to chance. Although the households were randomly assigned to neighboring housing units, there were only 27 pairs of units interviewed for each questionnaire version. More detailed analysis of the residency patterns of persons listed as "not usually living" in the unit provide more direct support for this conclusion. Of the 24 individuals listed in this category across questionnaire versions, 6 were identified as at high risk of being missed when using the standard SIPP household roster procedure.

While limited by sample size, therefore, these empirical results indicate that refining the household roster questions using a simple residency rule should improve coverage of individuals typically missed using the current household roster questions. The magnitude of the improvement in coverage cannot be estimated from these data for at least three reasons. First, the sample frame and sample sizes for these data do not allow for estimating the improvement. Second, it is not clear what rules would be used to define eligible populations. And third, even once rules were adopted it is not clear whether adopting a specific rule would hurt, rather than improve, overall coverage of the US population. While a residency rule may improve comprehension errors, it may increase other types of errors associated with deriving estimates of population parameters.

##### **4.1 Incorporating Residency into Respondent Definitions**

There a number of response and estimation issues that would have to be resolved prior to incorporating the alternative procedure into the normal SIPP definitions. First, eligibility criteria would have to be developed that allowed concentration on household members who are contributors to the financial status of the family/household. Clearly individuals who are just visiting the unit for a short period of time are not of primary interest. Alternatively, individuals who consistently spend some time in the household are of interest. Definitions based on previous residency (e.g., number of days), current residency (e.g., living here now?) and future residency (e.g., number of days in the future) may be used to develop a more acceptable definition. One might even want to incorporate other variables directly related to the income of the unit (e.g., payment of rent, contribution to other household budget items).

Second, it would be important to accurately calculate the probability of selection for a particular respondent. If an individual lives in two different locations, for example, then he/she may have more than one chance of falling into the sample. One way to solve this would be to use a multiplicity estimator based on the number of days the individual stayed in the household. The latter would be obtained by either asking the household respondent or by directly asking the household member.

A third, related, issue would be to decide on appropriate follow-up rules and methods for developing longitudinal weights. Longitudinal weighting is currently a topic of research both within and outside the Bureau and it is unclear how one would solve this problem. Intuitively, however, it seems that if one moved towards eligibility based on length of residence, then it may be possible to develop weighting schemes based on this criteria.

Whether a better estimate of the population would result from making these changes is an open question. By using residency rules, other types of response errors will be introduced. For example, respondents may miscalculate the number of days spent at a particular location or may even forget to report a particular location. This may introduce more bias into the population estimates than now exist. Similarly, respondents may not truthfully answer questions on residential patterns for many of the most important respondents in the household. Will better data result given these problems? The answer to this depends on the ability of future research to develop an appropriate compromise between the different types of errors.

#### **4.2 Future Research**

Future research could proceed in a number of different directions. The first would be to replicate the present study on a larger scale. The purpose of this project would be to better understand why results between the two versions differ and, more specifically, how respondent interpretation of the current SIPP items leads to underenumeration.

This replication would include:

- increasing the sample size,
- development of a more extensive set of debriefing items to assess the criteria used to define "usually lives here". This might include, for example, some type of detailed recall that allowed the interviewer to contrast different situations and probe why certain individuals were or were not included on the original roster.
- recruit and train interviewers more carefully. This would facilitate the use of a more complicated debriefing questionnaire and may enable more taping of the interviews.
- targeting households more specifically. This could include actually conducting a reverse record check study (e.g., unemployed, convicted criminals) or drawing sample from areas known to be underenumerated from the recent Census.

A second research design would be a more focussed investigation into how respondents are presently interpreting the current household roster questions on the SIPP. While the evidence in the literature clearly suggests that asking about "usual residence" is subject to different interpretations, it is not clear how much of the current undercoverage problem is due to this fact. The results from the present study suggest that a significant proportion could be avoided, at least when initially rostering the household, if a more explicit definition is used.

This could be completed in a cognitive laboratory setting by asking respondents to develop household rosters for different types of living situations (or vignettes) that varied by different parameters (e.g., length of stay, relationship to household members). Respondents would be recruited across different demographic, household and living situations. The results would be used to search for patterns of key criteria people use when making the decision to include or exclude individuals on the household roster. This approach has recently been used by BLS to better understand decisions individuals make when filling out the Consumer Expenditure Diary (Levin, et. al, 1991).

The disadvantage of this approach is that the artificial nature of the testing situation may not generalize completely to normal field conditions. In this case:

1. the use of "vignettes" may be too hypothetical for many respondents. That is, respondents who are not actually living in the situation described in the vignette may not use the same decision criteria as those individuals who are living in that situation,
2. the use of volunteers eliminates those respondents who would normally not provide a complete household roster, and
3. the testing situation increases motivation and recall to the point of changing the decision rules used by respondents

Notwithstanding these problems, we believe a study of this type would provide important information on the impact of the current questions. This is something that is still not very well understood at this point in time.

These two approaches could be combined by attaching the vignettes as debriefing questions to a replication of the field work proposed above.

A third direction future research could take would be to try to understand the living situations of the "extra" people listed on the household roster. Are they in households that are not covered by normal SIPP procedures? Would they be hard to follow and incorporate into the survey? Some of this research was conducted for the 1990 Census (Lowry, 1986) using information from the 1980 Census. This suggests that similar or more extensive analyses could (or are) being done using 1990 data. Original data collection would involve enumerating households using the alternative approach and then following all individuals not defined as "usually" living in the house to alternative addresses listed. The households at these addresses would then be enumerated. This would assist in understanding who these extra people are and might even provide more information on the type of criteria one would want to use when trying to operationally define individuals who are should be eligible for the SIPP interview.

A different approach to this issue would be to enumerate the household and follow-up the unit, perhaps over a 6-12 month period. Comparisons of the household composition at both points in time would provide information on the stability of the members defined using the new rostering approach.

In summary, the findings from the research reported above clearly support the need to assess logical ways to modify the current household roster questions contained on the SIPP. While clearly much of the coverage problem is due to respondents intentionally not listing individuals in the household, some portion is due to simple comprehension error (e.g, error associated with interpretation of "usual" residence). It is important for future research, therefore, to shift the burden of interpretation from the respondent to the designers/users of the SIPP by making the definition more explicit to the respondent. This research has suggested that one logical way to do this is to base the criteria on residence during a well defined reference period. It should be the goal of future research to verify that this is the case and how to adapt alternative questions in statistically tractable and meaningful ways.

## References

- Hainer, P., Hines, C., Martin, E. and G. Shapiro (1988) "Research on Improving Coverage in Household Surveys" Proceedings of the Fourth Annual Research Conference, pp. 513-539, U.S. Bureau of Census.
- Levin, K., Cantor, D., Lichtenstein, C., Brandt, S., Hosang, A., Stunkard, J. and K. Moesinger (1991) "Classifying Food Purchases on the Diary for the Consumer Expenditure Survey" Final Report submitted to the Bureau of Labor Statistics under Contract # J-9-J-8-0083.
- Lowry, Ira S. (1986) "Dealing with Residential Ambiguity in the 1990 Census of Population" Paper presented for the Census Conference on Residence Rules, December 15-16, 1986.
- Shapiro, G. and D. Kostanich (1988) "High Response Error and Poor Coverage are Severely Hurting the Value of Household Survey Data" Paper presented at the Annual Meetings of the American Statistical Association, 1988, Survey Research Methods.
- Wilson, W.J. (1987) The Truly Disadvantaged. Chicago: University of Chicago Press.



Exhibit 2. Household Enumeration Questions Based on Overnight Residence at Unit (Version 2)

FILL ITEMS 5 - 7 FOR ALL PERSONS LIVING OR STAYING HERE			
5 PERSON NUMBER	6a HOUSEHOLD ROSTER	6b RELATIONSHIP TO REFERENCE PERSON (RP)	7 PREVIOUS WEEK
Assign 201, 202, etc.  Circle person number for respondent	What are the names of all persons who have stayed here for at least one night since the first of July. Start with the name of the person or one of the persons who (owns, rents) this home.  Please include middle and maiden names.  (SEE CONTINUATION AT BOTTOM)  Last, First, Middle, Maiden	Ask if not apparent: What is ...'s relationship to... (Read name of RP)?  Example: Reference person, wife or husband, son, daughter, son-in-law, sister, foster son, partner, lodger, lodger's wife.	Did ... stay here in the previous week?  1 - YES 2 - NO  YES    NO
			1 <input type="checkbox"/> 2 <input type="checkbox"/>
			1 <input type="checkbox"/> 2 <input type="checkbox"/>
			1 <input type="checkbox"/> 2 <input type="checkbox"/>
			1 <input type="checkbox"/> 2 <input type="checkbox"/>
			1 <input type="checkbox"/> 2 <input type="checkbox"/>
			1 <input type="checkbox"/> 2 <input type="checkbox"/>
			1 <input type="checkbox"/> 2 <input type="checkbox"/>
			1 <input type="checkbox"/> 2 <input type="checkbox"/>
			1 <input type="checkbox"/> 2 <input type="checkbox"/>
			1 <input type="checkbox"/> 2 <input type="checkbox"/>
			1 <input type="checkbox"/> 2 <input type="checkbox"/>
			1 <input type="checkbox"/> 2 <input type="checkbox"/>

6a. (Cont).

- Am I missing anybody who has stayed here for only one night?
- Am I missing anybody who has visited for a short time, like a relative or a friend?
- Am I missing anybody who has stayed here because they temporarily needed a place to live?
- Is there anybody else who has stayed here in the last 30 days? Please remember that all the information you provide is strictly confidential and will not be used for any purpose, other than conducting this interview.
- Is there anyone who normally stays here but is away at school, on vacation or in the hospital?

If you wish, you may only provide the first name of the person.

Table 1. Demographic Distribution of Residents Listed on the Household Roster by Questionnaire Version\*

	Version 1 (Total (n=81) in Household)		Version 2 (Total (n=104) in Household)		
	<u>Frequency</u>	<u>Percent</u>	<u>Frequency</u>	<u>Percent</u>	
<u>Age:</u>					
17 - below	23	28%	19	18%	
18 - 25	9	11%	10	10%	
26 - 40	13	16%	28	27%	
41 - 60	18	22%	34	33%	
61 - higher	18	22%	13	12%	
<u>Gender:</u>					
Male	33	41%	53	51%	
Female	48	59%	51	49%	
<u>Race:</u>					
Black	77	95%	104	100%	
White	3	4%	--	--	
Asian	1	1%	--	--	
<u>Marital Status:</u>					
Married - Spouse Present	18	30%	28	32%	
Married - Spouse Absent	0	--	2	2%	
Widowed	7	12%	10	11%	
Divorced	10	16%	7	8%	
Separated	3	5%	10	11%	
Never Married	22	37%	30	34%	
<u>Education:</u>					
Never or Kind	1	2%	0	0%	
Elementary	7	12%	7	8%	
High School	37	62%	53	62%	
College	15	25%	25	29%	
<u>Age by Sex:</u>					
Male	17 & under	14	17%	11	11%
Female	17 & under	9	11%	8	8%
Male	18-25	4	5%	6	6%
Female	18-25	5	6%	4	4%
Male	26-40	5	6%	13	13%
Female	26-40	8	10%	15	14%
Male	41-60	4	5%	16	15%
Female	41-60	14	17%	18	17%
Male	61 & above	6	7%	7	7%
Female	61 & above	12	15%	6	6%

\*Includes all interviewed households.

Table 2. Demographic Distribution of Residents Listed on the Household Roster by Questionnaire Version\*

	Version 1 (Total (n=65) in Household)		Version 2 (Total (n=72) in Household)		
	<u>Frequency</u>	<u>Percent</u>	<u>Frequency</u>	<u>Percent</u>	
<u>Age:</u>					
17 - below	16	24%	13	18%	
18 - 25	6	9%	8	11%	
26 - 40	11	17%	18	25%	
41 - 60	16	24%	24	33%	
61 - higher	16	24%	9	12%	
<u>Gender:</u>					
Male	26	40%	33	46%	
Female	39	60%	39	54%	
<u>Race:</u>					
Black	61	94%	72	100%	
White	3	5%	--	--	
Asian	1	1%	--	--	
<u>Marital Status:</u>					
Married - Spouse Present	18	35%	20	33%	
Widowed	6	12%	9	15%	
Divorced	8	16%	7	11%	
Separated	2	3%	6	10%	
Never Married	17	33%	19	31%	
<u>Education:</u>					
Never or Kind	0	0%	0	0%	
Elementary	5	5%	4	8%	
High School	25	38%	36	50%	
College	12	18%	16	22%	
<u>Age by Sex:</u>					
Male	17 & under	10	15%	6	8%
Female	17 & under	6	9%	7	9%
Male	18-25	2	3%	6	8%
Female	18-25	4	6%	2	3%
Male	26-40	4	6%	7	9%
Female	26-40	7	11%	11	15%
Male	41-60	4	6%	10	14%
Female	41-60	12	18%	14	19%
Male	61 & above	6	9%	4	5%
Female	61 & above	10	15%	5	7%

\*Includes households where both members of the original pair were interviewed.

Table 3. Demographic and Residency Information for Persons Reported as Not Usually Living in the Unit by Questionnaire Version and Household

Household	Respondent	Sex	Age	Relation To Reference Person	Number of Nights in Unit	Still In Unit?	Stay Next 4 Months?	Reason for Stay
<b>VERSION 1</b>								
1	1	F	20	Friend's Daughter	1	NO	NO	Visit
2	1	M	17	Son's Friend	66	NO	NO	Visit
3	1	F	2	Granddaughter	5	YES	NO	Visit
	2	F	5	Granddaughter	5	YES	NO	Visit
	3	M	35	Son	5	YES	NO	Visit
4	1	M	35	Friend	1	YES	YES	Unclear
5	1	F	48	Cousin	--	--	--	--
6	1	F	62	Sister	3	NO	NO	Unclear
	2	F	60	Sister	3	NO	NO	Unclear
	3	M	63	Brother-in-Law	3	NO	NO	Unclear
<b>VERSION 2</b>								
1	1	F	35	Daughter	3	NO	NO	Visit
	2	F	38	Daughter	5	NO	NO	Visit
2	1	F	44	Cousin	1	NO	NO	Visit
	2	F	40	Cousin	3	NO	NO	Visit
	3	M	48	Brother	3	NO	NO	Visit
	4	M	49	Brother	3	NO	NO	Visit
	5	F	47	Friend	2	NO	NO	Visit
	6	F	39	Friend	1	NO	NO	Visit
	7	M	40's	Friend	1	NO	NO	Visit
	8	F	42	Friend	1	NO	NO	Visit
3	1	M	30	Husband	5	YES	DK	Not Specified
4	1	M	27	Brother	30	YES	DK	Move Out of Apt.
	2	F	47	Mother	7	NO	NO	Visit
5	1	M	18	Grandson	8	YES	YES	Not Specified
6*	1	F	8	Granddaughter	30	NO	NO	Visit

\*Taken from a household unit that did not have a matching completed case for Version 1.