

SOME DATA COLLECTION ISSUES FOR PANEL SURVEYS WITH APPLICATION
TO THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

Anne C. Jean and Edith K. McArthur, Bureau of the Census

Introduction: Need for a Longitudinal Survey.

The Survey of Income and Program Participation is designed to collect data which will improve our understanding of the income distribution, wealth, and poverty in this country. Information collected in the survey will be useful for planners and program administrators in areas such as income support programs and health care. The survey is longitudinal in the sense that the same persons are interviewed periodically over an approximately 2 1/2 year period. This implies following persons and updating information that reflects changes in their lives and in the composition of the households of which they are members--before, during, and after these changes occur. Persons in SIPP are interviewed every four months. At each interview, household members 15 years old or over are asked to report on income sources, amounts and employment for each of the previous four months.

With SIPP data we will be able to observe the effects over time of changes in receipt of different types of income upon the total income of a household; we will also see the effects of household composition change, such as the birth of a child or a marital separation, on participation in Federal transfer programs. In the past, analysts have often relied upon the income data collected in cross-sectional surveys, such as the March supplement of the Current Population Survey (CPS). The CPS describes household membership at a point in time, while obtaining income data for the entire previous calendar year. These data are consequently dependent on the household respondent's recall of events over the whole previous year. Thus many assumptions are made and monthly data cannot be collected accurately.

Implementation of a Longitudinal Survey.

The 1984 panel is the first panel of the SIPP. During the four months constituting Wave 1, that is October 1983 through January 1984, Census interviewers visited approximately 26,000 addresses located in 174 primary sampling units (PSUs) nationwide. The addresses were evenly distributed among four rotation groups, and each month one rotation group is assigned for interview. Nine interviews at four month intervals were scheduled for three rotations; the fourth rotation was scheduled for eight interviews.

The shift from an address sample for the first visit to a person sample in subsequent visits presented unique challenges to the planning staff, regional office staffs, and interviewers. Updating procedures for the address listings, noninterview classifications, interviewing procedures, and many other activities required for surveys maintaining an address sample were not appropriate. New controls and follow-up procedures, some requiring interregional office cooperation, were implemented. Interviewers received extensive training on new noninterview

classifications and movers' procedures. Office staff maintained extensive clerical controls to guarantee the receipt of control cards and questionnaires from interviewers and to monitor the processing of over 40,000 person records that were uniquely identified.

The remainder of this paper describes the Wave 1 field procedures associated with the implementation of the address sample and the follow-up procedures for subsequent waves. Included is an explanation of the SIPP identification system and those field operations designed specifically for sample maintenance and control. Some preliminary results of the 1984 panel follow-up are given and finally proposals for improving the follow-up system in future panels are discussed.

Wave 1 Address Sample Procedures.

Field activities for the first SIPP interview were similar to operations undertaken for other major surveys that are basically cross-sectional, such as CPS and the National Crime Survey (NCS). Interviewers listed specific addresses of living quarters either prior to or at the time of the interview visit. Reasons for differences between the number of expected units based on census address lists and the number of units listed by the interviewer were researched by the office staffs. During the first interview, the address was verified, the unit was classified as a housing unit or OTHER unit according to census definitions. Coverage questions were asked to determine if EXTRA or additional units were located at the address, and the interview status of the address was recorded.

The interview status distinguished interviewed households from noninterviews. Noninterviews were further classified by type. For example, Type A noninterviews include all eligible households for which interviews were not obtained, such as refusals or cases where no one was home each time the interviewer visited. Types B and C noninterviews were recorded for addresses containing no eligible household such as vacant addresses, or units under construction or being demolished.

In an interviewed household the interviewer listed all persons currently living or staying at the address, and applied a set of household membership rules to classify each person. Listed persons were classified as household members if the sample address was their usual place of residence as of the date of interview. The specific rules for household membership in SIPP are identical to those used in CPS. All household members listed in Wave 1 were designated as sample persons. After listing all household members, demographic information, such as age, sex, and relationship, was obtained for each household member and a SIPP questionnaire was completed for each household member who was 15 years old or over.

Development of Special Procedures for the Longitudinal Survey.

The procedural differences between SIPP and most other major household surveys conducted by the Census Bureau begin with the second interview. While other major surveys such as the CPS and NCS return to the same address for each subsequent visit regardless of whether the occupants of that address change, the SIPP interviewer returns to interview the same sample persons--that is, persons listed during the first interview. If persons move to a new address, they are followed and interviews are obtained at the new address. Between March 1981 and March 1982 almost 17 percent of the population of the United States moved.^{1/} If SIPP did not follow movers from the original sample households, we would lose the capability of observing the effects of many major changes in the original sample households, and the person sample would be biased, since it would not include movers.

Interviewers who discover that a Wave 1 sample person has moved (usually while updating the household roster) are instructed to inquire for new addresses at the original address and if further inquiry is necessary they are to contact mail carriers, rental agents, real estate companies and postal supervisors. Other sources may be used, such as an employer or a contact person (this is a person identified by the respondent during the initial interview as one who would usually know where the respondent was, such as a relative or a close personal friend). Occasionally, interviewers contact other persons with the same last name listed in local telephone books, although this procedure is not specified in their follow-up instructions. Beginning with the third interview visit, change of address notification forms are left with respondents and respondents are encouraged to mail these to the census regional offices (the address of the appropriate census regional office is preprinted on the form). In addition, advance letters are mailed to respondents before each interview; if the respondent no longer lives at that address the post office is requested to provide a forwarding address.

The regional office staff determines whether a new address will be assigned for a personal visit. Personal visits are required for all new addresses located in SIPP PSUs or within 100 miles of a SIPP PSU. Telephone interviews are encouraged for all sample persons who have moved to an address located more than 100 miles of a SIPP PSU within the United States. The following persons are excluded from mover follow-up:

- (1) Persons who join Wave 1 sample persons in later waves are not followed to new addresses unless these additional persons remain with Wave 1 sample persons who are 15 years old or over;
- (2) Persons who move out of the sample universe are not followed. These are persons who become institutionalized, move outside of the United States or live in an Armed Forces barracks;
- (3) Children under 15 who move and are not accompanied by a sample person who is 15

years old or over are not followed. The geographic area covered by personal visit follow-up is extensive. Based upon the 1980 census population distribution, about 130 million persons live in areas within SIPP PSUs; another 87 million persons live within 100 miles of the outer boundary of a SIPP PSU. We counted 226 million persons in 1980; 217 million are within our currently covered areas--96 percent of the population.

Of the 17 percent of the population that moved between March 1981 and March 1982, the great majority moved only a short distance--about two-thirds of the movers stayed in the same county (10 percent of the total population).^{2/} If persons or households move within the same county (or to a nearby county), the new address is usually assigned to the same interviewer for follow-up. The remaining third who move outside of their original county are usually transferred to another interviewer. This occasionally involves a transfer between two census regional offices.^{3/}

Of the 26,024 addresses included in the original SIPP address sample, 19,878 addresses were interviewed households in Wave 1 and were reassigned for a second visit. The 6,146 addresses reported as noninterview at the time of the first visit were not reassigned. Of these noninterviews, 1,019 were eligible households whose members refused to participate in the survey, or were temporarily absent, unable to be located or not interviewed for other reasons. Survey planners were reluctant to reassign in Wave 2 those Wave 1 eligible noninterviews because of the added complexity for both the interviewers and the processing system.^{4/}

Interviewers visited sample addresses for the second interview during February through May 1984 and attempted to locate and interview the approximately 40,000 sample persons interviewed during the first visit. New persons not present initially were added to the household rosters, provided original sample persons were still included on the roster. Any new persons who were household members 15 years old or older were also eligible for interview. If no sample person remained at an address, no interviews were conducted at that address, but interviewers were required to follow the sample persons to their new addresses.

The SIPP Identification System.

The SIPP Identification System is a numbering system designed to provide a unique unchanging identifier for each person in an interviewed household. The person identifier is used to link data from more than one interview for the same individual regardless of what moves have taken place or what changes in household membership have occurred since Wave 1. In addition, the ID system provides the means for grouping individuals into unique households in each wave. This is an important attribute, which allows for the tracking and identification of changing household membership--persons moving away can be linked to each household of which they have been a member since their first interview. However, no attempt is made during the field operations

to define or number each "different" household for longitudinal analysis.

The components of the operational SIPP identification system are:

- PSU number - 3 digits
- Segment number - 4 digits
- Serial number - 2 digits
- Address I.D. - 2 digits
- Entry address I.D. - 2 digits
- Person number - 3 digits

The PSU and segment numbers are assigned by Washington staff during sample selection. The 3-digit PSU number identifies a county or group of counties and is the same number used by other census surveys, such as the CPS and the NCS. As a sample of segments, that is, clusters of housing units, is drawn from a PSU, the segments are uniquely numbered within each PSU, using a 4-digit number. The clusters generally range in size from two to four housing units. Office staff in the 12 regional offices are responsible for assigning the 2-digit serial number. The 2-digit serial number is assigned sequentially in Wave 1 to each SIPP living quarters within a segment.

The 9-digit combination PSU, segment, and serial number uniquely identifies each sample address for the first interview. As a result, SIPP households interviewed during Wave 1, (October 1983-January 1984) can be uniquely identified with these three components: PSU, segment, and serial number. The PSU, segment, and serial numbers never change, regardless of movers and new household formations.

For SIPP, a 2-digit address ID code is added by office staff to provide a means for identifying more than one unique household associated with the same PSU, segment, and serial number. This situation occurs after Wave 1, when an original Wave 1 household splits up to form more than one household. The first digit of the address I.D. code indicates the wave a new address is first assigned for interview. The second digit sequentially numbers households originating from the same PSU, segment, and serial number. While not essential for Wave 1, an address ID code of 11 was assigned to all Wave 1 sample addresses. In later Waves, as SIPP sample persons move to new addresses, the office staff assigns new address ID codes to each new address brought into the survey by movers. Address ID codes assigned during a previous wave are deleted from the processing system for the current and successive waves if no SIPP sample persons remain at the address. Thus, the combination of PSU, segment, serial number, and address ID code uniquely identifies each sample address at each given Wave. As only one sample household is associated with a sample address, this combination provides unique household identifiers for a given Wave.

The person identification number is a 5-digit number consisting of an entry address ID code and a person number. It is assigned by the interviewer as each person is initially listed on the household roster. As the interviewer lists the name of each person in the household, he/she

transcribes the current 2-digit address ID code to each person's record. The 2-digit number is the entry address ID. Next, the interviewer assigns a 3-digit person number to each person. Numbers 101, 102, and so on, are assigned to persons at the sample address in Wave 1; the numbers 201, 202, and so on, are assigned to persons added to the roster in Wave 2; and so forth. The first digit indicates the wave the person enters the survey. This 5-digit number consisting of entry address ID and person number is not changed or updated, except in rare instances of merged households which are described later.

Thus, the 14-digit combination of PSU, segment, serial, entry address ID, and person number uniquely identifies each person in the SIPP survey and can be used to link data for persons across waves. The PSU, segment, serial, and address ID code uniquely identifies each household in a given wave; and the PSU, segment, and serial number can link all households in subsequent waves back to the original Wave 1 household.

An example of the numbering scheme may help to clarify it further. Consider a Wave 1 household. There is a basic control number consisting of PSU, segment and serial number, along with the address ID code. At the time of the first visit, four persons are listed--a father, mother, son and daughter. Each is assigned the current address ID code 11, along with a three digit person number--101, 102, 103, and 104.

The interviewer returns four months later and finds that the mother and father remain at the original address. The two children have moved to separate new addresses and both have married. The separate new addresses retain the basic control number (PSU, segment and serial number). One new address receives address ID code 21, the other receives 22. A new person, the son's wife is added. She is added at an address coded 21 in Wave 2, so she receives an entry address ID of 21 and person number 201. The daughter's husband is added at an address coded 22, so his person ID is 22-201. The original persons, the son and daughter, do not change their person ID's.

In Wave 3, the mother and father retire and move to Florida. No one lives at the original address. The mother and father moved in Wave 3, so their new address ID code is 31. Their person ID's remain the same. The son and his wife haven't moved in Wave 3. Their address ID's do not change. The daughter is still at the same address, so her address ID doesn't change. However, she has split-up with her husband and he has moved out. Since her husband is not an original Wave 1 sample person, he is not followed to his new address.

As mentioned previously, the operational phase makes no attempt to apply longitudinal household definitions to the changing relationships, nor to number households longitudinally. However, as analysts develop longitudinal definitions, the current data base must be able to provide the information required to support these definitions. Further refinements in the questions asked at each interview may be implemented as the needs of

a longitudinal household definition become more precisely specified.

The SIPP numbering system has several advantages over alternative schemes that have been considered:

- (1) The portion of the control number consisting of PSU, segment, and serial number is similar to the numbering system used by other major surveys conducted by the Bureau.
- (2) Interviewers are able to assign person numbers during the course of the interview. The person number is used in various parts of the questionnaire during the interview. This number is also transcribed to several other survey documents during the interview and immediately afterward during clerical coding operations. A person number assigned after the time of interview does not provide this immediate linkage.
- (3) The person number itself has relatively few digits, reducing the possibility of transcription errors.

Several disadvantages have been noted:

- (1) Duplications of person numbers for additional persons (persons added after Wave 1) can conceivably occur in situations where households have split and are in different regional office jurisdictions. The computer processing system identifies these duplicates and the regional office staff corrects them during processing.
- (2) Mergers between two separate sample households require special procedures. If this situation occurs, one set of controls is retained for the merged household. New person numbers are assigned to those persons who lose their original identifiers. Interviewers record both the old and new ID numbers on the control card to provide a means for linking the two ID's. By the end of the second Wave, this had occurred once.

Monthly Cross-Sectional Households.

While the ID system provides identifiers for each household in a given wave, it does not identify households for a given month. Monthly cross-sectional households are not constructed in the field; rather they are constructed during processing using information obtained during each wave. During each visit, demographic characteristics such as changes in marital status, changes in reference person (householder) status, and changes in household relationships are recorded on a control card. The same control card is used for each visit to the same address. If a sample person moves to a new address, the interviewer prepares a new control card for the new address and transcribes any information that is not expected to change. Date entered (month and day) and date left (month and day) are recorded on the control card for every entry and exit from an address. Reasons for entries and exits are coded:

Entry 1 - birth)
 2 - marriage)

3 - other)
4 - 5/)
Exit 5 - deceased
 6 - institutionalized
 7 - living in Armed Forces
 barracks
 8 - moved outside of country
 9 - separation or divorce
 10 - person who joined a household
 during Wave 2 or later and who
 is no longer living with any
 sample person
 11 - other
 99 - listed in error

Date entered and left is used during processing to group persons into households for a given month. A person entering a household before mid-month is considered to be a member for the entire month; a person entering after mid-month is considered not to be a household member for that month. A similar mid-month cutoff date is used for persons leaving households. As this monthly household determination is done during processing, it does not affect field operations, short of obtaining month and day of entries and exits.

Clerical Field Controls.

The SIPP movers' procedures have long been recognized as ambitious, requiring a system of field controls that are more extensive than those in effect for other major surveys conducted by the Bureau. Two standard forms are used for controlling interviewer assignments, and a third control was developed specifically for SIPP. All three forms are used during a clerical check-in at the regional offices.

An interviewer's Assignment and Control form is completed for each interviewer, listing every case in a given interviewer's assignment. A copy is sent to the interviewer and a control copy is kept in the office. As completed questionnaires are returned to the office, they are checked in against this form. A second control form lists all interviewers and the number of assigned cases for each interviewer. Tallies are kept as material is returned. This form gives supervisors a summary of the number of outstanding cases for a given month. The third control developed specifically for SIPP is a computer-generated listing of all persons listed as household members in Wave 1. It includes names, person numbers, interviewer codes, and interview status. The regional offices update the listing during each wave and account for every interviewed person as documents are received from interviewers. These three forms provide the basis of the clerical check-in and control. They must be updated to account for assignments that are transferred between interviewers and between regional offices, and they must be updated to include new persons entering SIPP after Wave 1. Two other control forms are used by the offices to facilitate the movers' operation. One form is used to list the original address of a sample household along with all subsequent addresses. It is used primarily to control the assignment of address ID codes. A second form, a worksheet, is used for transferring cases from one interviewer to another by telephone. Because

of time constraints, transfers are done by telephone; and required control card information-- such as new address, names of persons, demographic information for the movers, etc.--must be obtained from the original interviewer and passed on to the new interviewer.

While the scope of this paper concerns field operations, some mention must be made of two major features in the computerized processing system designed for check-in and control.

- (1) During the keying operation all persons listed on the control card who are 15 or over and are current household members must have an accompanying questionnaire. This check is done automatically at each keying station. Keying is done in the regional offices and immediate resolution of missing questionnaires is required.
- (2) At the end of each of the four months of each wave, a centralized check-in is completed in Washington. A control card record must be transmitted for every person showing an active status on a master file maintained in Washington of all active records. Offices cannot close out an interview month until every active status person is accounted for and some demographic data--age, race and sex--is verified to make sure that we are not checking in the wrong person. Each missing case is referred to the appropriate regional office for resolution.

Experience with Following Movers.

Available data for follow-up interviews conducted during February-May 1984 gives an initial indication of the success rates for the SIPP follow-up.

- (1) Percentage of movers found: about 80%.
- (2) Percentage of movers lost: about 20%-- represents 0.9 percent of all eligible SIPP households.

When sample persons move from the address at which they were contacted in the previous Wave (four months before), interviewers are instructed to go through a series of steps to locate the new address. If all the steps are "Dead ends" they fill in a form which describes what they know about the mover situation for those sample persons. A review of the forms for Wave 2 available at the time this paper was written (they are submitted on a flow basis and a form was not submitted for some of the cases) illustrated the kinds of events that took place leading to the sample person's moving without leaving a trace. In about half of the cases all household members moved leaving no forwarding address. For another quarter of the cases one or more persons had left the household leaving other members behind but those other persons had no information about the departee's whereabouts. In an additional 15 percent of the cases, the spouse (usually the husband) left the rest of the family and the remaining spouse could not or would not give a forwarding address. The remaining cases showed a variety of events; for example the person had moved and had no permanent new address, rather he was just

staying with various friends but the interviewer had no success in contacting him. The interviewers' comments showed considerable efforts in attempting to track these movers.

Recommendations for future SIPP Panels.

Improvements in the processing system and the expansion of follow-up procedures are envisioned for future panels. These recommended changes are intended to improve sample coverage in a number of areas.

In the 1984 panel, persons who leave the sample universe--become institutionalized, leave the country, or live in an Armed Forces barracks--are dropped from the sample.^{6/} As of the 1980 census, about 2.5 million persons were currently inmates of institutions such as mental hospitals, homes for the aged and correctional institutions. Another 613,000 persons were living in military barracks. Demographers estimate that about 160,000 persons emigrate from the United States each year.^{7/} As average stays in nursing homes are less than 60 days and live discharges account for about 75 percent of the discharges, a sample person who goes into a nursing home is likely to come out before the end of the SIPP panel. According to current procedures, members of each of these groups are reinstated only if they rejoin a SIPP household.

For the SIPP panel beginning in January 1985, planning is underway to track sample persons who become institutionalized. Interviewers will obtain the name of the institution in which the person is residing. At each subsequent interview they will determine whether the person is still there and if the person has been discharged they will obtain a new address. It will then be possible to follow sample persons leaving institutions even if they do not rejoin active SIPP households. There are no current plans to track sample persons who move outside of the country or to an Armed Forces barracks.

Interviewers may return to an address in the 1984 panel and find that all original Wave 1 sample persons have left but one or more additional persons (who joined households with sample persons after Wave 1) remain. In the 1984 panel no interviews are conducted at that address even though persons currently at the address lived with sample persons during at least part of the reference period. For future panels a final interview will be conducted for the additional persons remaining at the address. As in the 1984 panel, no subsequent follow-up is planned for these persons.

As described earlier, in the 1984 SIPP, only persons who are 15 or over are followed to new addresses; sample persons who are under 15 years old are not followed unless they move with a sample person who is 15 or over. However once they become 15 they are eligible for interview along with other members of their households. They are missed in the 1984 panel if they move before turning 15 and are not accompanied by a sample person who is 15 years old or older. Their absence may result in some bias in the survey data. In future SIPP panels, all sample persons who are 12 years old or older at the time of the

first interview will be eligible for follow-up. When a person who was 12 years old at the time of the first interview moves by him- or herself to a new address, occupants of the new household will be interviewed according to standard procedures--that is persons 15 years old and over will be administered a questionnaire. When the sample person turns 15, that person will also be administered a questionnaire.

A number of other recommendations have been made for future SIPP panels. These include:

- (1) Reassigning Wave 1 eligible noninterviews in Wave 2. Interviewers will be provided with instructions for obtaining household rosters and assigning person numbers retrospectively--i.e., as of a date approximately four months prior to the date of the second interview.
- (2) Adjusting the computerized check-in system to allow for new serial numbers (representing persons or addresses) to be introduced in Wave 2. This will provide flexibility for including missed Wave 1 housing units.
- (3) Developing a questionnaire that is appropriate for telephone interviews. This could be administered to persons who are not followed for a personal visit.
- (4) Increased automation over the next few years will eliminate much of the time consuming clerical operations associated with the check-in, control and monitoring of assignments.

In summary, SIPP has attempted an ambitious undertaking by implementing and attempting to improve an extensive follow-up program. Data users will be the ultimate beneficiaries and judges of the program's success.

1/ U.S. Bureau of the Census, "Geographical Mobility: March 1981 to March 1982." Current Population Reports, Series P-20, No. 384. Issued February 1984, U.S.G.P.O.

2/ U.S. Bureau of the Census, op. cit.

3/ The United States is administratively divided into 12 geographic areas. Each area consists of a group of states under the jurisdiction of a census regional office.

4/ Wave 2 interviews for households not originally interviewed in Wave 1 require special procedures for constructing household rosters. For example, interviewers would need to obtain the names of persons living at the address as of a reference date four months prior to the Wave 2 interview. An appropriate Wave 1 person number would be assigned (see the SIPP Identification System explained later in this paper). However the 1984 computerized check-in system was designed to reject any Wave 1 person number that appeared for the first time in later waves.

5/ Code 4 is used in circumstances where a sample person moves to an address already occupied by persons not previously in SIPP. The persons not previously in SIPP are added to the roster and are coded "4."

6/ It was decided, not to obtain proxy information for sample persons (as well as other members of a household that has at least one resident sample person) who die while they are in a SIPP panel.

7/ Robert Warren and Jennifer Marks Peck, "Foreign-Born Emigration from the United States: 1960 to 1970," Demography, Vol. 17, No. 1 (February 1980), pp. 71-84.