THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

Tracking Persons Over Time

No. 23

Anne C. Jean and Edith K. McArthur

U.S. Census Bureau

January 1987

U.S. Department of Commerce U.S. CENSUS BUREAU

THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

TRACKING PERSONS OVER TIME

No. 8701 23

Anne C. Jean and Edith K. McArthur

SIPP Working Paper #23

January 1987

U.S. Department of Commerce BUREAU OF THE CENSUS

Survey of Income and Program Participation

TRACKING PERSONS OVER TIME

No. 8701

by

Anne C. Jean and Edith K. McArthur

January 1987

ACKNOWLEDGEMENTS

The authors wish to thank the reviewers of this paper for their helpful comments and suggestions, in particular Evan Davey, Chet Bowie and Don Fischer of the Demographic Surveys Division, and Daniel Kasprzyk and Don Dahmann of the Population Division.

TABLE OF CONTENTS

Pages

PREFACE

INTRODUCTION..... 1

PROCESSING CONTROLS AND ACCOUNTS, A TWO-TIERED SYSTEM..... 3
Local Controls and the Growth of Automation..... 3
Centralized Controls..... 4

MOVERS, A CHALLENGE TO THE SYSTEM..... 6
Evolution of Movers' Rules in the 1984, 1985, and 1986 Panels..... 6

SUCCESSES AND FAILURES, TRACKING OUTCOMES..... 8
Characteristics of Movers Compared to Non-Movers..... 10
Characteristics of Interviewed Movers Compared to Movers Missing the 5th Wave..... 11
Tracking Movers in Other Longitudinal Surveys..... 12

FURTHER RESEARCH ON MIGRATION..... 12

FUTURE DIRECTIONS FOR SIPP..... 13
Evaluating the Gaps..... 13
Improving Current Field Procedures..... 13
Processing and Control..... 13

CONCLUSION..... 14

Footnotes
References
Tables

PREFACE

The Survey of Income and Program Participation (SIPP) is a large scale panel survey with a rotating design. New panels are introduced each year, and interviewed over a 2 1/2 year period. Persons interviewed in the first wave are tracked throughout the life of the panel and are interviewed at 4-month intervals.

This paper describes the tracking rules and the control system which work together in SIPP to ensure the receipt of properly identified survey records for each eligible person in the SIPP sample.

INTRODUCTION

The Survey of Income and Program Participation (SIPP) is a nationwide survey designed to provide detailed information on the economic condition of persons, families and households in the United States. The SIPP collects information on cash and noncash income, participation in government transfer programs, labor force status, assets and liabilities, and other topics which affect persons' economic well-being. The collection of this kind of information is unique and is instrumental in studies of national issues such as tax reform, Social Security program costs, and health insurance coverage. For example, the effects of changes in eligibility rules or benefit levels on different demographic groups can be determined, or the results of alternative taxing arrangements can be observed. The primary goals in implementing the SIPP were to improve reporting of income and other program-related data and to do so in a way that would allow for analysis of changes over time at various levels: person, family, household, etc.

The SIPP is one of only a handful of longitudinal surveys which track persons. Most surveys are cross-sectional, based upon an address sample, and any return visits are made to the address, regardless of the residents. In contrast, the SIPP starts a panel with an address sample but then the individuals residing at those addresses at the time of the first interview form the sample. They are called sample persons, in the jargon of the SIPP. For the 2 1/2 years that a panel is in the field it is those sample person who must be followed. Sample persons who move are interviewed at their new addresses and any other persons who live with them are interviewed also as long as they remain living with sample persons.

Approximately 16.8 percent of the U.S. population moved in the 12 month period from March 1983 to March 1984 (CPS, 1986). Therefore if that trend held and if movers were not followed, by the time a panel ends, after 2 1/2 years, a large percentage of the original sample membership would no longer be interviewed. Therefore, as with other longitudinal surveys, it was recognized that movers need to be followed as long as they remain in sample.

A mover is defined as any sample person who changes physical address. Thus a move can be across counties but within the same metropolitan area, across state lines, across the country, or even out of the country. The tracking system for the SIPP was designed to accommodate any of those types of moves. For example, the same interviewer might conduct an interview if the new address is within the same assignment area, or a different interviewer working out of the same regional office might be assigned the sample household, or the interview for that household might be conducted over the telephone or out of another census regional office.

After a developmental program consisting of a site research test and two experimental national panels in the late 1970's, SIPP fielded its first national panel in October 1983. This is referred to as the 1984 Panel. New panels were then introduced in February of each year, beginning in February 1985. With the exception of the 1984 Panel, the duration of interviewing for each panel is approximately 2-1/2 years (31 months). The 1984 Panel extended over 34 months.

Households in each panel were assigned to one of four groups, called rotation groups, with followup interviews scheduled at four month intervals. Survey

information was obtained for the four months preceding the month of interview. The four month period was the usual reference period throughout the panel for most questions. Rotation group 1 in the 1984 Panel was first interviewed in October 1983 with a reference period of June-September 1983; rotation group 2 was first interviewed in November 1983 with a reference period of July-October 1983; rotation group 3 was first interviewed in December 1983; and finally rotation group 4 was first interviewed in January 1984. In February 1984, interviewers returned for a second visit to rotation group 1 households and asked questions covering the period October-January. As the 1984 Panel continued its cycle of nine interviews, the 1985 Panel began its own cycle of eight interviews in February 1985, and as both the 1984 and 1985 Panels continued their cyclic patterns, the 1986 Panel began in February 1986. With the introduction of the 1986 Panel, a three panel overlap occurs in several months of each calendar year.[1/] (See Nelson, et al, 1985.)

The initial SIPP interview establishes a listing of persons who are household members at the time of the interview. Subsequent interviews attempt to include these same persons along with any new persons living with them. The persons included in SIPP for the first interview are named sample persons and those who join in later interviews are named additional persons. Since SIPP attempts multiple interviews with the same persons over a period of time, survey planners designed a numbering scheme, field tracking system and data processing system which focused on the sample persons and additional persons as the basic units of observation. A system was designed to provide a unique unchanging[2/] identifier for each person (See Jean and McArthur, 1984), a comprehensive set of tracking rules and a check-in procedure which defined the universe of expected person records, verified the receipt of these records, and provided timely feedback to field offices to resolve any instances of missing person records. The tracking system currently encompasses a broad array of activities and involves hundreds of Census Bureau personnel throughout the United States: field interviewers who search for movers' new addresses, staff in the regional offices who maintain local controls, coordinate assignment transfers and provide supervisory support to interviewers, and Washington staff who are responsible for the design, maintenance and monitoring of the centralized control system, including analyses of tracking results. It is not a small task, as SIPP includes over 100,000 sample persons during those periods of 3-panel overlap (which covers about half of every year). Adding to the challenges of scale is a decentralized organizational structure with day-to-day survey operations under the direct control of 12 different administrative regions covering the United States.

The remainder of this paper describes the procedures used to maintain the sample, especially with regard to migration and provides some information about our success thus far. It describes the control system which is used to account for each individual sample member, and it contrasts the characteristics of movers with those of nonmovers. Lastly the paper provides a brief description of what two other longitudinal surveys are doing in this area, and presents some future directions SIPP.

PROCESSING CONTROLS AND ACCOUNTS, A TWO-TIERED SYSTEM

The control system for SIPP operates on two levels: a local level designed for individual regional office control of cases assigned to interviewers working in each of the 12 administrative areas and a centralized control in Washington which coordinates the activities of all 12 regional offices. It is an interactive system in which regional office staff assume a pivotal role in the two-way flow of information from field interviewers to regional offices to Washington and back.

INTERVIEWERS ⇄ 12 REGIONAL OFFICES ⇄ Washington

Local Controls and the Growth of Automation

When the 1984 SIPP Panel was introduced, each of the 12 regional offices was provided with identical instructions and forms for maintaining clerical control over the sample. SIPP used two standardized report forms commonly used on other Census surveys to control the sample at the household level, along with a computer generated Master Field Control unique to SIPP, which controlled the sample at the person level. The Master Field Control was a set of printouts generated by Washington after processing data for the first interview.[3/] Each regional office received a Master Field Control listing the cases within its jurisdiction. It contained each sample person's unique identifier, name, interviewer code (a code identifying the interviewer who completed the initial interview), and household interview status (a code indicating whether a household was interviewed or if not interviewed indicating the reason for the noninterview). Columns were set up on the Master Field Control for clerks to update this information for each successive interview. As completed assignments were returned from interviewers, clerks in the regional offices updated household interview status, added additional persons to the list, recorded any changes in interviewers' assignments and assigned a code to movers, called the address ID code. Properly completed, the Master Field Control was used as a regional administrative tool to ensure the receipt of survey materials for every expected sample person across all interview periods of a given panel. After regional office check-in and clerical editing, the office staff keyed the survey documents, performed a set of simple edits using the data entry equipment, and then transmitted the data to Washington for centralized check-in and comprehensive editing.

During 1984, microcomputers were installed in all regional offices and the clerically maintained standardized report forms were gradually replaced by a relational database program using R:base 4000. The program was used to develop assignment lists, produce mailing labels, check in completed work, and produce a number of administrative reports. This program was upgraded to R:base 5000 during the early months of 1986 and continually improved to provide supervisors with a more powerful and efficient management tool. Offices maintained parallel clerical controls during the early stages of automation until the new system was operating smoothly and personnel were adequately trained. The timing of the elimination of parallel clerical controls continues to vary by regional office. By mid-1986 about half of

the regional offices had stopped using the standardized clerical forms. By January 1, 1987 all regional offices will be fully automated, dropping all clerical controls. At that time, an automated Master Field Control will also be in place, and the system will be upgraded once more to R:base System V.

Centralized Controls

Prior to data acceptance for the second interview, centralized Master Control Files (MCF) are created in Washington. These files extract data submitted from the first interview from all 12 regional offices. A separate MCF is created for each rotation within each panel. As the sample reaches three concurrent panels, this requires 12 separate Master Control Files, with three files in use for any month containing all three panels. It also requires that the 12 files be updated for each new wave.

Each MCF contains two record types, a person record and an address record. (See Attachment A.) Each sample person and additional person, including children, has his/her own person record. The person record contains the person's name, unique person identifier and other identification codes such as a regional office code and an address ID code. The address ID code is a two digit number assigned by regional office staff to each new mover address. All household members at a given address are associated with the same address ID code for a particular interview since they live at the same address. In addition to these codes, demographic information such as age, sex, race, and marital status are included. Status codes are assigned during processing, to indicate whether a person record is required, not required but accepted if received, or not expected. For example, if a person became institutionalized and thus out-of-scope, an inactive status code would be assigned. This indicates that a person record is not expected.

The address record contains a limited amount of information for each address ID code ever assigned. It includes an unchanging unit control number, household interview status for each wave of interviewing, regional office code for each wave and status code. The status code indicates whether the address is active or inactive. Inactive addresses are no longer visited for SIPP. An address can become inactive for a number of reasons, such as all sample persons leaving the address, reductions in sample size, or noninterviews not assigned for future interviews.

Data keyed in the regional offices are transmitted to Washington and subjected to a pre-edit consisting of a match to the MCF and a consistency edit. A match to the MCF is required for every incoming unit control number. (The control number is an unchanging identifier assigned to the original sample unit. Households that move keep the same control number.) After matching the control number, a series of checks are performed on the incoming address ID code for that control number. Returning address ID codes requires a match, while new mover address ID codes are added to the file. Next, matches at the person level are done, requiring a match to the MCF on unique person identifier, age, sex and race for each person. The program also identifies new additional persons, verifies their identifiers, and adds them to the MCF.

Occasionally, during the check-in, we discover that information on the MCF is incorrect. Only Washington staff are authorized to make changes to demographic information on the MCF based on new information obtained by

interviewers. Sex and age corrections are required more often than others. (See Kalton, McMillen and Kasprzyk, 1986.) Thus, cross-sectional files that are linked, but not longitudinally edited, will show some discrepancies in these basic characteristics.

Movers are controlled by a series of checks which basically require two records for each mover eligible for followup. One record indicates that a person has left an address by including one of several "left" codes indicating the reason for leaving along with a date left. A second record indicates that the mover has moved into a new address by giving one of several "entered" codes with the reason for entering along with a date entered. A series of checks screens eligible movers from ineligible movers. For example, a record with left code 05 "deceased" obviously would not require a second record with an entered code. (See Attachment B for a list of entered and left codes.)

Cases not passing the check-in stage of processing are electronically transmitted back to the regional offices, where reject listings are printed out. Clerks research the errors by referring to control cards and questionnaires, and occasionally contacting interviewers. Then corrected listings are keyed and electronically transmitted back to Washington. After passing all the matches required for check-in and correcting any errors, cases progress to a series of consistency edits. Errors in consistency edits are also corrected following the same process.

A few days before close-out for a rotation, a list of missing persons is generated. A missing person is a MCF person record with a status indicating that a record is expected but not received. All missing person cases and rejects require resolution before any new rotation is processed. This check-in and reject resolution process originally caused considerable delays and backlogs of work awaiting processing. In fact, there was concern that the introduction of a third panel in 1986 would be the proverbial straw that broke the camel's back. However, since the introduction of the 1986 panel, processing deadlines in almost every month have been met. These deadlines are usually set for 3-4 weeks after the last day of an interview month.

This brief description presents a broad view of a very complex process. As the 1984 Panel began, and as the data poured into our system, we wondered if it would support the complexity and dynamism of our mobile and ever changing population, as people are born, become married, change their names, divorce, join the Armed Services, split up and come back together and continue changing. Secondly, we wondered if the system would be too complex or too demanding. Could we realistically expect data to pass through a very detailed and rigorous pre-edit and not become hopelessly entangled or lost in the system? Or given the large scale of SIPP, was it too optimistic to expect an accounting for every eligible sample person with no tolerance for a small percentage of missing cases? As the 1984 Panel progressed, we found that the control system worked well. It guaranteed the receipt of survey documents for every eligible person and did not place excessive burdens on field staff. [4/]

Even so, the need for some modifications became apparent after several interviewing cycles. For example, we improved our ability to keep track of children under 15; entered and left codes were extended to identify a wider variety of reasons for persons entering and leaving; special entered and left codes were developed for a person added to an address who should have been

added in a previous wave, as well as for those who left in a prior wave. We found that several cases which were checked in and later deleted from the data file were not retransmitted and the MCF still considered the case checked-in, due to the earlier transmittal. The program was modified to re-set the MCF status for cases deleted from the data file after check-in. Despite the need for some modifications as we learned more about SIPP's requirements, the processing system performed very well and supported the various field procedures and tracking rules. These procedures and tracking rules also underwent modifications as SIPP progressed from October 1983 until the present.

MOVERS, A CHALLENGE TO THE SYSTEM

The movers' rules adopted for the SIPP required a careful balancing of theoretical considerations with the operational factors of cost, complexity for the field staff, and allocation of computer programmers' time for designing and redesigning the programs required for data processing. The following is a chronological account of the development of the movers' rules from the 1984 panel through the 1986 panel, followed by a summary of tracking outcomes.

Evolution of Movers' Rules in the 1984, 1985 and 1986 Panels

The movers' rules established at the start of the 1984 panel specified that all sample persons (persons included in SIPP for the first wave of interviewing) who were fifteen years old or older were eligible for followup. This included children who became 15 years old during the panel. However they only become eligible for followup after they became 15 years old. There were certain restrictions to the general rule. Any movers to places outside of the sample universe were not followed. The sample universe included the noninstitutional population of the United States, excluding members of the Armed Forces living in barracks. For persons who moved outside of the sample universe, interviewers simply assigned one of three codes to the mover, indicating whether the person had been institutionalized, moved out of the country, or was living in a barracks. Interviewers obtained a limited amount of additional information, such as the name and address of the institution, but this additional information was not keyed or processed further.

Additional persons who joined SIPP households after the first interview remained in sample as long as they lived with sample persons. Thus if these additional persons moved and no longer lived with sample persons, they were not followed. Since each interview collects data covering the previous four months, not obtaining interviews for these additional persons results in missing data for any months within the four month reference period when the additional persons were still living with sample persons. (However the date of separation from sample persons is recorded.) We did not attempt to obtain the missing information by interviewing a proxy.

Finally, sample persons who moved within the country but more than 100 miles from a SIPP PSU (Primary Sampling Unit area, usually one or several contiguous counties) were not followed for a personal visit. Attempts were made to retain these persons in sample by conducting telephone interviews. Most movers, however, did remain within the personal visit followup areas since only about 4% of the U.S. population lived in areas beyond the 100-mile limit.

Some further restrictions were placed on the 1984 panel noninterview followup. If an eligible household was not interviewed during the first wave, it was not reassigned for interview during subsequent interview periods. However, there was some flexibility in adding sample persons who were missed in interviewed households. If a sample person was inadvertently excluded when a household was interviewed during the first wave, the missed person could be added to sample at the time of the second interview. One drawback to the 1984 procedures was that the person's unique identification number had to correspond to a numbering range reserved for additional persons. The 1984 panel processing system would not accept any wave 1 person identifiers appearing for the first time after the wave 1 interview. Thus, tracking was not assured if the missed person moved in later waves, as interviewers would be likely to identify the person as an additional person, and not an original sample person.

A procedural committee met prior to the start of the 1985 panel to consider a series of recommendations to modify the tracking rules and adjust the processing system. Several of the recommendations were not implemented because they required extensive reprogramming and testing, a commitment of staff time and funds that could not be met with available resources. However, some changes were adopted for the 1985 panel.

Eligible households that were not interviewed during the first wave were reassigned during the second interview period. This occurred in the 1985 panel during June – September of 1985. Missed sample persons who were added during the second interview were assigned a unique identifier in the range reserved for wave 1 sample persons. The processing system was adjusted to accept these cases during the second interview. A special entered code was also used so the new procedure would not weaken the controls on the sample.

Beginning in May 1985 for the continuing 1984 panel and in October 1985 for the new 1985 panel, we began tracking institutionalized sample persons. Interviewers verified the status of institutionalized sample persons by inquiring at a sample household or by telephoning the institution. If the sample person left the institution, interviewers attempted to obtain a new address and conduct an interview if the person was no longer institutionalized.

The procedural committee's recommendations that were not implemented were:

(1) Lower the age for following sample person movers to 12 years old. This would have ensured the inclusion of children who become 15 during the panel. Currently, these children remain in sample only if they continue to live with a sample person who is 15 or over.

(2) Obtain final interviews from all additional persons who no longer live with sample persons. This recommendation was limited to addresses where the original sample persons had moved out, but the additional persons remained. This limitation was placed on the recommendation to simplify field procedures.

(3) Obtain a final proxy interview for persons who leave the sample universe. This recommendation did not include deceased persons. We had planned to include only data covering those months when the person was still in sample.

Two other proposals were postponed for future consideration, but have since been dropped. These were:

(1) Obtain final interviews for additional persons who move and no longer live with a sample person.

(2) Include in the mover followup, all additional persons who were not in the sample universe as of wave 1.

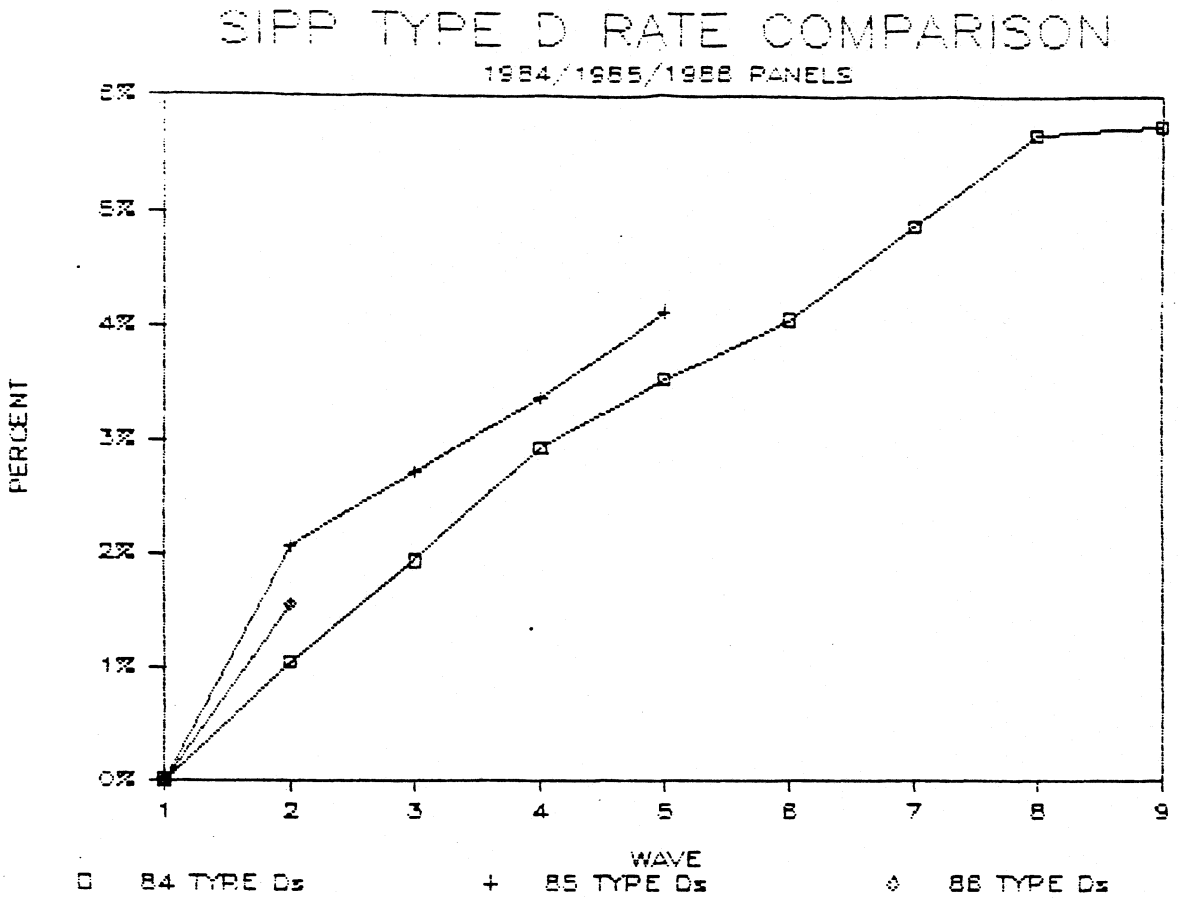
With the introduction of the 1986 panel, no new or expanded procedures were implemented. Rather the practice of reassigning eligible wave 1 noninterview households for followup at the second interview was discontinued. Processing wave 1 returning noninterviews in the 1985 panel had proved more burdensome than originally anticipated, and relatively few cases were brought back into the sample. In addition, any eligible household missing two consecutive interviews was not contacted again.[5/]

No changes have been proposed for the 1987 panel. Thus, it is likely that the procedures currently in place will be continued.

SUCCESSES AND FAILURES, TRACKING OUTCOMES

Rates at which households were "lost" from sample because a new mover address could not be found are closely monitored. These households are included in a classification named Type D noninterviews and are presented in the graph below for the 1984, 1985 and 1986 Panels. For purposes of defining a Type D household, any sample person who moves and is living at an unknown new address is considered to be a member of a Type D household. Thus, if one person moves and cannot be found, while others remain and are interviewed, we count one Type D noninterview household and one interviewed household. If an entire household moves and cannot be found, we count one Type D noninterview household.

The Type D rates shown on the graph are cumulative. Most households that become a Type D noninterview, remain so in subsequent interviews (unless we find them). New Type D households are added to the total number of households. At the end of the 1984 Panel we had lost close to 6% of eligible SIPP households due to unfound movers. By the end of wave 5 of the 1985 Panel we had lost slightly over 4%.



This graph is taken from the SIPP Operational Reports Series, Report # 5

Another way to look at how well we track movers and keep them in sample is by looking at persons instead of households. We looked at 5 waves of the 1984 panel for this analysis. Movers had to be identified using a combination of three variables on the cross-sectional file. (For a description of this process see Robbin, et al, 1986.) Table 1 shows the results. Of our total sample (movers and non-movers combined) 79 percent completed all 5 interviews, 17 percent did not complete the fifth interview, and 4 percent were missing one or more interviews but were interviewed during the fifth wave. Among persons who moved sometime after the first interview, that is between waves 2 and 5, 69 percent completed all 5 interviews, 23 percent did not complete the fifth interview, and 9 percent were interviewed in the 5th wave but were missing at least one intervening interview.

Next we examined movers by their interview status as of the fifth wave. The results are shown in Table 2. Approximately 11 percent of all persons who moved between waves 2 and 5 ended up being classified as having moved to an unknown address. This represented about 2.24% of our total number of persons in the study.

Characteristics of Movers Compared to Non-Movers

Studies of movers and of sample attrition in SIPP are being used to evaluate tracking rules; they form the basis for recommending changes in field procedures; and will be used to evaluate current procedures for making noninterview adjustments. Since it appeared that we were not able to keep movers in sample as well as we could the non-movers, we looked at the characteristics of the movers.

This section describes the characteristics of sample persons who moved during the first 5 interviews of the 1984 panel. Included in the analysis are only those sample persons 15 years old and over eligible for all 5 interviews.[6/] We found that 20 percent of our sample moved at least once after their Wave 1 interview in the 16 months from Wave 2 through Wave 5—a rate fairly representative of the U.S. population as a whole.

We hypothesized that movers would be different from non-movers. Table 3 compares the characteristics at the time of the first interview of non-movers to movers. Chi-square tests, including a sample design effect were performed on the distributions of characteristics of the two groups shown in columns 2 and 3. In the table, a one in parenthesis beside a characteristic indicates that the distribution of that characteristic at the time of the first interview was different between movers and non-movers.

Some of the mover characteristics that we found to be significant might change as a result of a move, in the sense that the move was the occasion that provided the context for the change. These characteristics include: type of housing, urban or rural residence, and area of residence. In fact, persons liable to move for other reasons may have selected their residential characteristics during a previous move in order to facilitate later moves. Movers' housing was different—movers tended to rent, non-movers to own their homes. Movers were more likely to have lived in a metropolitan area than were non-movers. At the beginning of the panel there were proportionately more persons living in areas under the administration of the Atlanta and Dallas regional offices—two areas experiencing substantial growth recently—who would move during the period between the second and the fifth interview. The states surrounding the Boston, New York, and Philadelphia regional offices, in contrast, had proportionately more non-movers.

Other characteristics recorded during the first interview, such as age, sex, race and educational attainment may condition persons to be more amenable to undertaking a move. Movers tended to be younger than non-movers; a much higher proportion of movers were between 15 and 44 years old than non-movers. Proportionately more movers were of Spanish origin than non-movers. Movers have higher educational attainment than non-movers. Movers were less likely to have a savings account or other types of assets than were non-movers. However there were no significant differences in the distributions of sex and race between movers and non-movers.

Another group of characteristics may be related to a move. That is, the move may have been precipitated by a change in one or more of these characteristics or the move may have been precipitated by a desired change in these characteristics. The distribution by household size was significantly different between movers and non-movers. Movers tended more often to be

classified as "child of reference person," other relative, or non-relatives. Movers were more likely to be never married than non-movers, and non-movers tended more often to be married than the movers. Movers were more likely to be employed than non-movers. They were more likely to be receiving cash or non-cash benefits, such as AFDC or food stamps. Also, movers' household monthly income and personal monthly income tended to be lower than that of non-movers.

Characteristics of Interviewed Movers Compared To Movers Missing the 5th Wave

We compared characteristics of movers that are successfully interviewed through 5 waves with mover who are not interviewed in the fifth wave. Chi-square tests were performed to determine whether significant differences in the distributions of the characteristics shown in columns 4 and 5 of Table 1 were present. Characteristics that we found were different between these two groups are designated with a "2" in parenthesis by the characteristic. Of interest here is that some of the variables which were significant among movers and non-movers are also significantly different between movers that we follow through five interviews and those that apparently attrit; these include whether or not persons are of Spanish origin, whether they owned or rented their home, marital status, educational attainment, employment status, whether or not they lived in a metropolitan area, household monthly income, and asset ownership. In addition, distributions of race for the two groups were significantly different.

The last two columns in Table 1 show persons who were members of the apparent attrition group classified by the reason recorded for their non-interview. Only the two major reasons are shown, household refusal and moved to unknown address. Two points must be considered in looking at these columns: firstly, the fact that the person moved may have been unrelated to the household refusal, and secondly, as these columns contain persons who moved at least once, the designation of "moved to unknown address" may have been recorded after one or more moves were successfully tracked. Again, chi-square tests were performed to determine whether the distributions of characteristics for the two attrition groups were significantly different. Those that were found to be significantly different are designated with a "3."

The distributions of marital status, asset ownership, and receipt of cash and non-cash benefits were found to be significantly different.

The above discussion should not be interpreted as implying a measured change in the characteristics described; the characteristics shown were as recorded in the first interview and may have remained unchanged throughout the five interviewing waves for movers as well as non-movers. For example the recorded marital status in all five interview waves may have remained "married, spouse present" regardless of whether the person's residence changed during the period; or similarly a move may have taken place but the person's relationship to reference person may have remained "child of reference person." As the analysis of the potential interrelationships of moving to various other events has not been undertaken yet, the above discussion simply indicates the kind of study that can be undertaken with the SIPP. One of the principal purposes of the SIPP is to track changes in persons' lives, and to determine the interrelationships between events which appear to occur concurrently—such as a job change and a move, a marital status change and a move, and so on. The analysis of these potential relationships is a next step which needs to be undertaken.

Tracking Movers in other Longitudinal Surveys

Other longitudinal surveys follow similar procedures to track movers and maintain their sample. One such survey is the Panel Study of Income Dynamics (PSID) conducted by the Institute for Social Research at the University of Michigan. The survey which commenced in 1968 has been able to locate between 96 to 98 percent of its family units sample each year. (Each new family unit formed by sample members or their children increases the base on which this percentage is calculated.) The PSID pays its respondents, one per family unit, $10.00 for a completed interview. A further $5.00 is paid to the respondent if a postcard intended to determine whether an address correction is necessary is returned. Interviews take place once a year and the interviewing period runs from approximately March through October. About 90 percent of the interviews are conducted on the telephone. Procedures used to locate new addresses for movers are similar to those used in the SIPP: contact persons are obtained each year, and interviewers are instructed to contact other family members, utility companies, the telephone company, neighbors and so on, to obtain new telephone numbers and addresses.[7]

PSID had a higher proportion of movers than the SIPP or the CPS in 1983: 21.4 percent of the family units weighted and 24.3 percent unweighted. Possibly this higher proportion is due to the original sample selection which included oversampling of people who were poor in 1968 (see Institute of Social Research, 1985)

The Australian Longitudinal Survey (ALS) is a survey of the labor market experience of persons 15 to 24 years old. The survey began in 1984 with a list of 3,000 persons of which 2,400 respondents were contacted in 1984 and 1,900 responded in 1985. Unlike the SIPP, among nonrespondents the refusal category is second to the category of "address unknown." Mobility was found to be the most significant variable related to nonresponse in 1984. Respondents are interviewed once a year. In order to track respondents, the respondents are asked at the time of interview for names, addresses and phone numbers of 3 close relatives or friends who would be likely to know their whereabouts in the next year. Staff of the ALS keep contact with respondents twice a year between surveys; once with a Christmas card and once with a newsletter. The letter also contains a request for respondents to inform the ALS if they have moved (with a postage paid card). There were hundreds of responses to the mailouts. Contacting the contact persons was also found to be useful. Using the electoral rolls, however, did not add very much and probably was not worth the clerical effort. (For more information on the ALS see McRae, 1986).

FURTHER RESEARCH ON MIGRATION

One of the major goals of SIPP, as a longitudinal study, is to track short-term changes in persons' lives, such as a job change or occupation change or a change in marital status, and the correlation of the events to other events such as migration. Another area of interest is to look at the characteristics of persons' residences before and after a move-region of residence, state, whether living quarters are owned or rented (Dahmann, 1986). We strongly recommend that on the longitudinal files currently being developed at the Census Bureau a single recode be created which identifies persons who have moved while in sample.

FUTURE DIRECTIONS FOR SIPP

There are three critical areas, each of which is important to the continued success of the SIPP longitudinal panel. First, we must maintain an adequate set of tracking rules. In a purely theoretical milieu, these rules would ensure that all movers who remain members of the universe of eligible persons are retained in sample. When practical considerations of cost and feasibility are taken into account, we fall short of the theoretical ideal. Secondly, field procedures must fully support the tracking rules. This encompasses issues of special training, incentives, and techniques, and requires constant monitoring and evaluation. And thirdly, controls and processing systems must keep track of each person throughout all of the processing stages.

Evaluating the Gaps

As mentioned previously, several proposals to increase mover coverage have been considered, but have not been implemented. These include: following sample person movers who are 12 years old or older (instead of the current rule of 15 years); locating additional persons who no longer live with sample persons and obtaining final interviews; obtaining a final proxy interview for persons who leave the sample universe; including in the mover followup any additional persons who were not in the sample universe for wave 1; and reassigning wave 1 noninterviews. Our survey plans include an evaluation to determine the frequency of cases falling into these categories and an analysis of the effects of including and excluding these persons.

Improving Current Field Procedures

Current training on movers' procedures emphasizes technical aspects of the interviewer's job, such as proper coding of entered and left reasons and dates; when to complete a control card for a Type D household noninterview; and completing a control card at a mover's new address. More attention in training will be given to the task of discovering the new address, e.g., how to use contact persons and all other available sources of information. One improvement in the 1987 panel will be the recording of additional contact persons' names and telephone numbers. The additional contact persons are especially necessary when unrelated persons are living together. In current panels, only one contact person per household is required.

While Type D noninterview rates are monitored and field staff are encouraged to make every effort to find movers' new addresses, these rates are not used in current performance evaluations. Nor are there any supervisory guidelines establishing "unacceptable" or "needs improvement" rates. The bureau is studying the issue of incentives for finding movers, along with supervisory guidelines for acceptable or unacceptable loss rates due to unfound movers. These issues, along with the need to improve training on following techniques, have been the subject of current discussions at the bureau and were topics at a recent supervisors' conference.

Processing and Control

The control system for SIPP is more extensive than most surveys conducted by the Census Bureau. The results of matching cross-sectional data files is one measure of the control system's effectiveness which will continue. Since the

cross-sectional data files do not always have a clear and unambiguous accounting for each sample person, the work on matching cross-sectional files at the bureau, along with documentation of nonmatches, is essential.

Some questions remain on the amount of data processed for institutionalized persons. If a person leaves sample to become institutionalized, we currently do not process information on type of institution. Furthermore, once a person is out of the sample universe, no further information is processed for that person, unless he/she returns to sample. For example, if an institutionalized person dies we currently do not process information that the person has died. Interviewers simply make a note of the situation and cease any further inquiries about the person's status.

CONCLUSION

In summary, we believe that the enormous undertaking of tracking persons in SIPP, and accounting for every person in a systematic and consistent way has provided users of SIPP data with an invaluable database, contributing to a better understanding of how changes over time affect our economic well-being.

Footnotes

1/ The duration of the 3-panel overlap in calendar year 1986 is 6 months. In future years, the overlap may be slightly longer due to elimination of "short" waves. A "short" wave contains less than 4 months and occurred in Waves 2 and 9 of the 1984 Panel, Wave 2 of the 1985 Panel, and Wave 3 of the 1986 Panel.

2/ Person identifiers were changed in a few instances. These changes are documented for all user files. See Jean and McArthur, 1984 for further background discussing conditions under which identifiers are changed.

3/ The initial interview period also required both local and centralized check-in systems at the housing unit level. These check-in systems are not discussed in this paper.

4/ Creation of cross-sectional data files required the deletion of persons who left the sample universe before the 15th of the first month of the reference period. (The reference period is the four month interval preceding the interview month.) Thus, matched cross-sectional files will show instances of persons dropping out of sample with no apparent explanation. Documentation of these cases is provided to users of the cross-sectional files.

5/ Reasons for these changes are documented in an internal Census Bureau Memorandum For: Charles D. Jones, From: Gary M. Shapiro, Subject: SIPP 1986 - Treatment of Wave 1 Noninterviews and Sample Adjustments, January 8, 1986. The memorandum recognized the value of reassigning wave 1 noninterviews for reducing the risk of bias in cross-sectional estimates but expressed reservations concerning our ability to develop a missing wave imputation system for wave 1. It recommended that we not reassign Wave 1 noninterviews primarily to reduce the added burden on the programming staff.

6/ One rotation group, one quarter of the sample, was not scheduled for the second wave interview and thus over the time covered was eligible for 4 interviews not 5.

7/ Per conversation with Priscilla Hilebrandt of the Institute for Social Research on November 3, 1986.

References

Jean, Anne C., and Edith McArthur
(1984) "Some Data Collection Issues for Panel Surveys with Application to the Survey of Income and Program Participation," SIPP, Working Paper Series, No. 8407, "Papers Presented in the SIPP Session IV at the Annual Meeting of the American Statistical Association, Washington, D.C.: U.S. Bureau of the Census.

McRae, Ian
(1986) "Collection and Analysis of Longitudinal Labor Market Data" Presented to the Eighth Austrailian Statistical Conference, Adelaide, Australia.

Survey Research Center, Institution for Social Research, The University of Michigan
(1985) A Panel Study of Income Dynamics; Procedures and Tape Codes; 1983 Interviewing Year; Wave XVI; A Supplement, Ann Arbor, Michigan.

Kalton, Graham, and James Lepkowski
(1985) "Following Rules in SIPP," Journal of Economic and Social Measurement, Volume 13, Numbers 3 and 4, pp. 319-329.

Robbin, Alice, Martin David and Thomas S. Flory
(1986) "Facilitating Complex Data Management Tasks: An Example Using the Survey of Income and Program Participation for Studying the Determinants and Consequences of Migration Behavior." CDE Working Paper 86-18, Center for Demography and Ecology University of Wisconsin - Madison.

Dahmann, Donald
(1986) "Geographical Mobility Research with Panel Data," Growth and Change, Volume 17, July, pp. 35-48.

Nelson, Dawn, David McMillen and Daniel Kasprzyk
(1985) "An Overview of the Survey of Income and Program Participation: Update 1" SIPP, Working Paper Series, No. 8401, Washington D.C., U.S. Bureau of the Census.

U.S. Bureau of the Census
(1986) "Geographic Mobility: March 1983 to March 1984," Current Population Reports, Series p. 20, No. 407, Washington D.C: U.S. GPO,

Kalton, Graham, David McMillen and Daniel Kasprzyk
(1986) "Nonsampling Error Issues in the Survey of Income and Program Participation (SIPP)," in Proceedings of the Second Annual Research Conference, March 23-26, 1986, Washington D.C.: U.S. Bureau of the Census, pp. 147-164.

TABLES

Description of column headings for tables

(1) Universe—All 100 level persons 15 years old and over who were in rotation groups 1, 2, or 3 who remained after the sample reduction which took place around Wave 5.

(2) Non-movers—Persons whose address ID did not change, who were not Type D 24 (moved, address unknown) and who remained in the same state throughout the five interview waves.

(3) Movers—Persons whose address ID changed, or who were Type D 24, or who changed state of residence sometime during the five interview waves.

(Included in this column but not shown separately are 488 persons who were movers but they either were interviewed in the first and the fifth waves and were missing at least one of the interviews in between or they left the SIPP universe—due to death, institutionalization, an overseas move, or a move onto armed forces barracks.)

(4) Movers who stay in SIPP—Persons who, even though a move was recorded (as defined above), were interviewed by self or proxy in all five waves.

(5) Movers who leave SIPP—Persons who were movers and who were not interviewed in at least the fifth wave.

(6) Refusals—Movers whose recorded reason for noninterview was a household refusal.

(7) Type D 24's—Movers whose recorded reason for noninterview was a Type D 24 (moved, address unknown).

Characteristics shown are as recorded at the first interview. Beside each characteristics is a note in parentheses. These refer to the results of three statistical tests which were done for each of the characteristics, such as sex, race, age, etc., to determine whether the persons represented in a column were significantly different from another column shown in the table.

If significant differences were present between columns 2 and 3 - Non-movers versus Movers - the parenthesis contains a "1." If significant differences were present between columns 4 and 5 - Movers who stay in SIPP versus Movers who leave SIPP - the parenthesis contains a "2." And if significant differences were present between columns 6 and 7 - Movers who refused and Movers who were Type D 24 - the parenthesis contains a "3." If none of the tests revealed a significant difference for a particular characteristic the parenthesis contains only a dash ("-").

Table 1: Comparison of Interview Completion by Mover/Non-Mover Status Across Five Interviews (Rotation groups 1,2,3; Persons 15+)

	Total		Movers		Non-movers	
	#	%	#	%	#	%
Total	25,138	100.0	5,069	20.2	20,069	79.8
		100.0		100.0		100.0
All five interviews	19,878	79.1	3,485	68.8	16,393	81.7
Missing at least the fifth interview	4,222	16.8	1,148	22.6	3,074	15.3
With fifth interview but missing at least one	1,038	4.1	436	8.6	602	3.0

Table 2: Movers' Interview Status at Fifth Interview

	Number	Percent
Total Movers	5,069	100.0
Movers interviewed all five waves	3,485	68.8
Movers who missed one+ interviews but interviewed in fifth wave	436	8.6
Household refusals*	86	1.7
Moved to unknown address*	137	2.7
Movers, out of universe*	3	-
Other**	210	4.1
Movers missing at least fifth interview	1,148	22.6
Household refusals*	350	6.9
Moved to unknown address*	564	11.1
Movers, out of universe*	52	1.0
Other**	182	3.6

* Status recorded for first missed interview.

** Includes temporarily absent, no one home, etc.

TABLE 3: Characteristics of Persons by Mover/Non-Mover Status: SIPP 1984 Panel
(After 5 Interviews, Persons 15 Years and Over, Rotations 1,2,3)

	Universe (1)	Non-Movers (2)	Total Movers (3)	Movers who stay (4)	Movers who leave SIPP (5)	Refusals (6)	Type D's (7)
Total Number	25,138	20,069	5,069	3,485	1,096	350	564
Wave 1 Variables							
RegionalOffice: (1,2)							
Boston	7.2	7.4	6.1	6.7	4.8	4.9	4.3
New York	6.9	7.3	5.3	3.1	11.5	6.9	15.4
Philadelphia	10.4	10.9	8.0	8.5	6.1	3.7	6.9
Detroit	8.4	8.6	7.5	8.1	4.3	6.9	2.7
Chicago	7.8	8.0	6.9	8.3	4.0	2.6	4.6
Kansas City	8.4	8.7	7.4	9.3	2.9	4.6	2.3
Seattle	8.6	8.3	9.6	10.5	7.2	10.0	6.2
Charlotte	8.9	9.0	8.6	9.7	6.4	4.3	6.4
Atlanta	11.2	11.0	12.0	9.6	19.0	26.3	16.1
Dallas	9.7	8.9	13.1	11.4	16.7	18.3	16.5
Denver	5.7	5.4	7.2	7.5	6.6	4.6	8.0
Los Angeles	6.8	6.5	8.1	7.1	10.5	7.1	10.6
Residence char:(1,2)							
Not an SMSA	25.5	26.0	23.2	25.4	16.2	16.6	14.4
SMSA: LT 100,000	1.3	1.2	1.8	2.0	1.7	4.0	0.0
SMSA:100-249 thou.	9.4	9.4	9.1	9.6	7.2	8.0	6.0
SMSA:250-499 thou.	9.2	9.3	8.8	9.6	7.4	9.1	6.7
SMSA:500-999 thou.	13.4	13.1	14.7	14.5	15.1	13.1	17.2
SMSA:1-2.9 mill.	24.1	23.2	27.3	25.3	32.6	37.1	31.6
SMSA:3-14.9 mill.	17.2	17.8	15.2	13.7	19.7	12.0	24.1
Living quarters: (1)							
House,apt.flat	94.0	94.5	91.9	91.2	93.7	94.9	94.0
Nontransient Hotel	0.2	0.1	0.4	0.4	0.6	0.0	1.1
Perm.in trans.Hotel	0.1	0.1	0.2	0.2	0.2	0.0	0.2
HU/Rooming House	0.1	0.1	0.1	0.1	0.2	0.3	0.2
Not HU/Rooming Hse.	0.1	0.1	0.1	0.1	0.1	0.0	0.2
Mobile Home,no add.	4.4	4.1	5.8	6.4	4.1	3.4	3.5
Mobile Home,w/add.	0.9	0.8	0.9	1.0	0.9	1.4	0.5
All other	0.3	0.3	0.6	0.7	0.2	0.0	0.4
Living quarters: (1,2							
Owned/Being bought	69.7	77.6	38.3	40.1	32.3	36.6	26.8
Rented for cash	28.1	20.2	59.1	57.0	65.5	61.1	71.3
Occ'd w/o cash pmt.	2.2	2.1	2.7	2.8	2.2	2.3	2.0
Race: (2)							
White	86.8	87.0	86.1	88.8	79.0	86.0	75.0
Black	10.4	10.3	10.9	8.8	15.8	8.3	19.9
Am.Ind/Esk/AlNative	0.4	0.4	0.6	0.4	1.0	0.3	1.2
Asian/Pac.Isl.	2.4	2.3	2.5	2.0	4.2	5.4	3.9

TABLE 3 - Continued

	Universe (1)	Non-Movers (2)	Total Movers (3)	Movers who stay (4)	Movers who leave SIPP (5)	Refusals (6)	Type D's (7)
Interview Status:(-)							
Self	67.0	66.8	67.8	68.6	66.8	65.4	63.7
Proxy	33.0	33.2	32.2	31.4	33.2	34.6	36.3
Interview Length: (1) (Minutes)							
Less than 15	27.4	27.3	27.7	27.5	27.7	28.3	27.3
15 to 29	43.8	43.2	45.8	46.2	44.8	42.0	44.9
30 to 44	20.9	21.1	19.9	19.8	20.7	22.9	20.6
45 to 59	6.0	6.2	5.4	5.3	5.5	5.4	5.7
60 or more	2.0	2.1	1.2	1.3	1.3	1.4	1.6
Number of Persons in Household: (1)							
1	11.6	11.5	11.9	11.8	12.7	11.1	14.7
2	29.0	29.6	26.6	26.9	25.5	25.7	27.1
3	20.3	19.8	22.2	22.4	22.2	24.9	20.6
4	20.0	20.3	18.8	19.7	17.1	17.4	15.4
5	10.7	10.7	10.8	11.1	9.6	8.9	9.4
6	4.4	4.4	4.4	3.2	6.6	5.1	6.9
7	2.1	1.9	2.5	2.2	3.6	5.1	3.4
8 or more persons	2.0	1.7	2.9	2.8	2.8	1.7	2.5
Sex: (-)							
Male	46.9	46.7	47.8	45.9	51.1	44.3	54.1
Female	53.1	53.3	52.2	54.1	48.9	55.7	45.9
Age: (1)							
15 - 24	21.9	18.0	37.2	37.2	36.8	30.6	40.2
25 - 44	37.7	35.4	46.9	46.5	49.9	48.9	50.2
45 - 64	25.2	28.8	11.1	11.3	10.7	15.1	8.3
65 and over	15.1	17.8	4.8	5.0	2.6	5.4	1.2
Ethnicity: (1,2)							
Spanish Origin	5.6	5.1	7.7	5.8	13.6	9.4	16.8
Not Spanish Origin	94.4	94.9	92.3	94.2	86.4	90.6	83.2
Relationship: (1)							
Reference Person	35.2	36.7	29.6	30.4	27.0	31.4	27.0
Primary Ind.	13.0	12.4	15.5	15.2	13.9	13.7	20.0
Spouse	28.5	30.2	21.8	23.0	18.1	25.7	13.8
Child	16.8	16.0	19.9	20.1	22.7	19.4	17.9
Other Relative	3.5	3.0	5.4	4.6	9.0	5.1	9.2
Non-rel w/rels.	0.4	0.3	1.1	0.8	1.7	0.6	2.7
Non-rel. no rels.	2.5	1.4	6.7	6.0	7.6	4.0	9.4

TABLE 3 - Continued

	Universe (1)	Non-Movers (2)	Total Movers (3)	Movers who stay (4)	Movers who leave SIPP (5)	Refusals (6)	Type D's (7)
Marital Stat: (1,2,3)							
Mar'd,spouse pres.	58.1	61.3	45.7	47.6	40.1	53.4	32.3
Mar'd,spouse absnt.	0.6	0.5	0.8	0.8	1.2	1.7	1.1
Widowed	7.3	8.3	3.5	3.6	2.6	3.4	2.5
Divorced	6.6	5.9	9.6	9.2	11.1	10.6	13.1
Separated	2.3	1.8	4.2	3.5	6.4	2.9	8.2
Never Married	25.0	22.2	36.1	35.2	38.5	28.0	42.9
High grade: (1,2)							
LE 8	11.4	12.4	7.5	6.3	8.6	6.3	9.2
9-11	16.8	16.8	17.0	14.7	22.3	17.1	27.8
12	35.8	35.8	35.9	36.1	35.9	41.4	32.3
GE 13	36.0	35.0	39.7	42.8	33.3	35.1	30.7
Employment: (1,2)							
With job:							
Worked all weeks	54.6	53.0	60.6	63.0	56.0	60.3	52.8
Missed 1+ weeks	1.2	1.1	1.4	1.4	1.3	1.1	1.4
Time on layoff	0.3	0.2	0.4	0.2	1.0	1.1	0.9
Job part of time:							
No layoff no looking	1.3	1.2	1.6	1.7	1.1	0.9	0.7
Did look or layoff	1.3	1.1	2.2	1.8	3.2	2.0	3.4
No job:							
All mo looked/layoff	4.6	3.8	7.7	6.3	11.7	7.1	15.6
Some looked/layoff	0.6	0.5	1.1	0.7	1.6	0.9	2.1
No looking/layoff	36.2	39.1	25.1	24.8	24.1	26.6	23.0
Hours Wk'd/Week: (1)							
Not applicable	37.1	39.5	27.7	26.1	30.5	27.4	33.7
1 to 19	5.7	5.8	5.1	5.9	3.2	3.4	3.2
20 to 34	8.9	8.4	10.8	10.3	10.9	10.9	10.8
35 to 40	33.3	31.9	38.7	39.4	37.9	42.3	34.9
41 or more	15.0	14.3	17.7	18.3	17.6	16.0	17.4

TABLE 3 - Continued

	Universe (1)	Non-Movers (2)	Total Movers (3)	Movers who stay (4)	Movers who leave SIPP (5)	Refusals (6)	Type D's (7)
Hhld.Mo.Income:(1,2)							
LE 299	4.3	3.9	6.1	5.4	8.2	6.0	10.6
300 to 599	7.7	7.4	9.0	7.9	11.4	5.7	14.5
600 to 899	8.2	8.0	8.9	8.7	9.2	9.1	9.4
900 to 1199	7.8	7.6	8.7	8.0	10.6	8.6	12.6
1200 to 1599	11.6	11.4	12.5	12.3	13.6	13.7	13.3
1600 to 1999	10.1	9.9	10.8	11.3	9.6	10.6	9.9
2000 to 2999	21.7	21.8	21.5	22.6	19.7	21.1	18.3
3000 to 3999	13.4	14.2	10.5	11.4	7.8	12.6	3.5
GE 4000	15.1	15.9	12.0	12.4	9.9	12.6	7.8
Person Mo.Income:(1)							
LE 299	30.3	29.8	32.6	31.3	36.4	31.7	40.4
300 to 599	16.0	16.2	15.3	14.5	16.1	11.7	18.8
600 to 899	12.5	12.2	13.5	13.9	12.1	14.9	9.8
900 to 1199	9.8	9.7	10.3	10.0	11.4	11.4	11.2
1200 to 1599	10.3	10.1	10.9	11.5	10.6	12.0	9.4
1600 to 1999	6.7	6.7	6.5	7.3	4.2	5.4	3.2
2000 to 2999	8.7	9.1	6.8	7.2	6.1	7.7	5.1
3000 to 3999	3.1	3.3	2.4	2.5	1.9	4.0	0.7
GE 4000	2.7	2.9	1.8	1.9	1.1	1.1	1.4
Asset Summary:							
Savings acct:(1,2,3)							
Yes	56.5	58.8	47.0	50.9	34.8	43.4	26.1
No	43.5	41.2	53.0	49.1	65.2	56.6	73.9
All other:(1,2,3)							
Yes	41.1	44.5	27.6	30.4	21.2	29.1	14.7
No	58.9	55.5	72.4	69.6	78.8	70.9	85.3
Hhld. recvs. cash benefits: (1,3)							
Yes	8.2	7.6	10.7	9.6	13.1	8.3	17.6
No	91.8	92.4	89.3	90.4	86.9	91.7	82.4
Hhld.recvs.noncash benefits: (1,3)							
Food stamps	7.2	6.2	11.0	10.0	13.2	5.4	18.6
Other only	9.8	9.3	11.5	11.2	12.8	9.7	14.0
No benefits	83.0	84.4	77.5	78.8	74.0	84.9	67.4

ILLUSTRATION OF A MASTER CONTROL FILE

ENTPER	TRANS	BATCH	DOC	PW-AGE	SEX	F-NAME	L-NAME
11101	0	0	0	27	1	ROBERT D	PUBLIC
11102	0	0	0	26	2	SUSAN	PUBLIC

All Persons Ever Processed At Control Number 999-1234-03

```
PSU = 999  SEGMENT = 1234  SERIAL = 03  ENTPER = 11101

ENTPER-REC-SER = 03
ENTPER-REC-ENTPER = 11101
AUID(WAVE) =  11 21 21 21 21 21  0  0  0
RO(WAVE) =    23 27 27 27 27 27  0  0  0
AGE-PREV = 027                        AGE-CURR = 000
SEX-PREV = 1                          SEX-CURR = 0
RACE-PREV = 1                         RACE-CURR = 0
MSTAT-PREV = 1                        MSTAT-CURR = 0
RRP-PREV = 01                         RRP-CURR = 00
PSTAT-PREV = 1                        PSTAT-CURR = 0
PNSP-PREV = 102                       PNSP-CURR =   0
PNPT-PREV = 999                       PNPT-CURR =   0
STATUS-PREV = 7                       STATUS-CURR = 7
WAVE-LEFT = 6
LEFT-CODE = 08
DATE-LEFT = 000501
PER-INT-STAT(WAVE) =  2  2  1  2  2  1  0  0  0
FIRST-NAME = ROBERT D
LAST-NAME = PUBLIC
TRANS# =  0     BATCH# =    0     DOC# =  0
```

Person Record

```
PSU = 999  SEGMENT = 1234  SERIAL =  3  ADID = 11

ADID REC SER =  3
ADID REC ADID = 11
HH-INT-STAT (9 WAVES) =  1 26  0  0  0  0  0  0  0
STATUS = 0
CURRENT-RO = 23
CCTRANS# =  0
CCBATCH# =   0
CCDOC# =  0

PSU = 999  SEGMENT = 1234  SERIAL =  3  ADID = 21

ADID REC SER =  3
ADID REC ADID = 21
HH-INT-STAT (9 WAVES) =  0  1  1  1  1 23  0  0  0
STATUS = 0
CURRENT-RO = 27
CCTRANS# =  0
CCBATCH# =   0
CCDOC# =  0
```

Address Records

ADID (WAVE) =	11 21 21 21 21 21 0 0 0
	The ADID shows which address ID code was assigned to the person's residence at each wave. All addresses in Wave 1 are assigned address ID code = 11. In this example the person moved at Wave 2, creating a new address ID code = 21. The person remained at 21 until Wave 6. At Wave 6 the person moved out-of-scope and is no longer at an eligible SIPP address. Therefore, zeroes are recorded in the 7th, 8th and 9th wave locations.
RO (WAVE) =	23 27 27 27 27 27 0 0 0
	Regional office 23 had the case in Wave 1. The new address in Wave 2 was under the jurisdiction of regional office 27. For Waves 7-9, no regional office is recorded, since the person was out-of-scope.
* AGE-PREV * SEX-PREV * RACE-PREV	These items carry the age, sex and race characteristics reported at the last wave. As new wave information is processed through the pre-edit, updated information is entered in AGE-CURR, SEX-CURR and RACE-CURR. Sex codes are 1 = Male, 2 = Female. Race codes are 1 = white, 2 = black, 3 = American Indian, Eskimo or Aleut and 4 = Asian or Pacific Islander.
MSTAT-PREV = 1 RRP-PREV = 01 * PSTAT-PREV = 1 PNSP-PREV = 102 PNPT-PREV = 999	These items show marital status, relationship to reference person, population status (code 1 if 15 yrs. old or older; code 2 if under 15), person number of spouse, and person number of parent (999 if no parent in household). PREV indicates that this is the previous wave information; CURR indicates new wave information. Marital status codes are: 1 = married, spouse present, 2 = married, spouse absent, 3 = widowed, 4 = divorced, 5 = separated and 6 = never married.

Illustration of a Master Control File Record

The example of a Master Control File Record is taken from an actual case record created for the 1984 SIPP panel, with any identifying information changed to protect the privacy of the individuals.

When the Master Control File is queried for a particular case, it produces a listing of all persons ever associated with a unique control number assigned to a Wave 1 sample unit. After Wave 1, this list could contain original sample persons along with any additional persons who joined them after Wave 1. The list could reflect multiple households after Wave 1, since persons move and split up, forming more than one household. This list is at the top of the illustration.

The unchanging identifier for the original sample unit consists of PSU = 999, SEGMENT = 1234, SERIAL = 03. The two persons listed at the original sample unit were assigned person identifiers 11101 and 11102. In this example, there were no additional persons joining the sample persons after Wave 1. The person record for 11102 (Susan Public) is not shown on the attachment. However, it is similar to that of 11101 (Robert D. Public), but with characteristics pertaining to 11102.

The original Wave 1 address, located in one of the states administered by regional office 23, is accounted for under the address record with ADID = 11. In Wave 2, both persons moved to a new address in a state administered by regional office 27, creating a new address record with ADID = 21. (Each regional office has its own 2-digit code which can range from 21-32.) In Wave 6, both persons moved out of the country. No further information is recorded after Wave 6, since both persons are out-of-scope at that point.

The following is an explanation for items on the person record and address ID records. An asterisk indicates that a match is required for this item during check-in.

* PSU = 999	- A three digit code assigned to a primary sampling unit which is made up of one or more contiguous counties.
* Segment = 1234	- A four digit code assigned to segments selected within the PSU. Segments consist of 2-4 addresses selected for sample.
* ENTPER-REC-SER = 03	- This is the serial number portion of the control number.
* ENTPER-REC-ENTPER = 11101	- This is the person identifier. When combined with the control number it provides a unique identifier.

STATUS-PREV = 7

This variable indicates whether or not a person record is expected. Codes are: 0 = record is expected, 6 = record is not required, but will be accepted if it comes in, 7 = record is not expected and if one comes in it will be rejected, 8 = case is inactivated due to a sample adjustment and no record is expected. The STATUS-CURR is set for incoming wave information.

WAVE-LEFT = 6

This shows the most recent wave where a left code or entered code had been assigned. Left and entered codes are assigned to show reasons for entering and leaving an address. These codes are shown on Attachment B.

LEFT-CODE = 8

This code (08) indicates that the sample person moved outside of the country. See Attachment B for a complete list.

DATE-LEFT = 000501

This gives the latest date (month/day) recorded for leaving or entering an address. The 1984 panel did not have a year value for this item, but subsequent panels also included year.

PER-INT-STAT (WAVE) =

2 2 1 2 2 1 0 0 0

This variable shows the interview status of the person by wave. Codes are: 1 = personal interview, 2 = proxy interview, 3 = type Z refusal, 4 = type Z other, 0 = no questionnaire. A type Z noninterview was a person who was not interviewed in a household that had one or more completed interviews for other household members. Code 0 was entered on all person records where no one at the household had been interviewed.

FIRST-NAME LAST NAME	These items were copied from the survey documents each wave. They were not used for matching during check-in.
* TRANS # * BATCH # * DOC #	These were cross-sectional processing numbers assigned to questionnaires and control cards during the check-in. These processing numbers were posted to the file during check-in. Records which were recycled during the correction process were required to maintain the same processing control number.

Each address visited for interview is assigned a address ID code. In the illustration, two addresses were visited, the original wave one address (ADID= 11) and the mover address (ADID = 21). Therefore, two address ID records are created.

For address ID = 11 HH-INT-STAT (9 WAVES)=	1 26 0 0 0 0 0 0 0
	This item records the household interview code. The original Wave 1 address was a completed interview (code 1). At Wave 2, all sample persons had moved (Code 26). From Waves 3-9, no further information was processed for address ID = 11. See Attachment C for a list of HH-INT-STAT codes.
STATUS = 0	A control card for address ID = 11 was not processed during Waves 7-9 so no status code was entered. For control cards still being processed, codes range from 20-24.
	Code 20 = active case with at least one person record
	Code 21 = active case with no person records. (Used in the 1985 Panel for reassignments of Wave 1 noninterviews.)
	Code 22 = inactive case with a person record. This would have been assigned after Wave 6 processing, indicating that no Wave 7 address record was expected.

Code 23 = inactive cases with no person records.

Code 24 - inactive due to a sample reduction

For address ID code 21, the entries in HH-INT-STAT (9 waves) are: 0 1 1 1 1 23 0 0 0

This indicates that no interview was conducted in Wave 1. (The address was first visited in Wave 2.) Interviews were conducted in Waves 2-5. In Wave 6, a code 23 "entire household out-of-scope was recorded.

ENTERED/LEFT CODES

The 1984 Panel began with only 11 codes for entered and left. These were expanded with each new SIPP Panel and by 1986 there were 33 codes. All three panels used the expanded coding system by 1986.

1984 codes

Entered

1- Birth
2- Marriage
3- Other
4- (Code assigned to additional persons already living at address where sample person moved in.)

Left

5- Deceased
6- Institutionalized
7- Living in Armed Forces Barracks
8- Moved outside of Country
9- Separation or divorce
10- Person number 201+ no longer living with sample person

Codes Added in 1985

12 - left code for mergers.
13 - re-entering sample after one or more waves of inactivity
24 - sample person added at second visit
99 - deleted, originally listed in error

Codes Added in 1986

Entered - This Wave

16 - from Institution
17 - from Armed Forces Barracks
18 - from Outside the Country
19 - due to Separation or Divorce

Entered - Should have been added in a previous wave

21 - Birth
22 - Marriage
23 - Other
36 - From Institution
37 - From Armed Forces Barracks
38 - From Outside the Country
39 - Due to Separation of Divorce

Left - Should have been deleted in a previous wave

25 - Deceased
26 - Institutionalized
27 - Living in Armed Forces Barracks
28 - Moved Outside of Country
29 - Separation or Divorce
30 - 201+ Person no Longer Living with Sample Person
31 - Other

CODES FOR HOUSEHOLD INTERVIEW STATUS

	01	– Interviewed
	02	– No one home
	03	– Temporarily Absent
	04	– Refused
(Wave 1 only)	05	– Unable to locate
	06	– Other Type A Noninterview
(Wave 1 only)	09	– Vacant
	10	– Occupied by persons with URE
	11	– Unfit or to be Demolished
	12	– Under construction, not ready
	13	– Converted to temporary business or storage
	14	– Unoccupied site for mobile home, trailer, or tent
	15	– Permit granted, construction not started
	16	– Other Type B
(Waves 2–9)	16	– Entire household institutionalized or temporarily ineligible
(Wave 1 only)	17	– Demolished
	18	– House or trailer moved
	19	– Converted to permanent business or storage
	20	– Merged
	21	– Condemned
	22	– Other Type C
(Waves 2–9)	22	– Deleted (sample adjustment, error)
	23	– Entire household deceased, moved out of country, or living in Armed Forces Barracks
	24	– Moved, address unknown
	25	– Moved within country beyond limit
	26	– All sample persons relisted on new Control Card