THE SURVEY OF INCOME AND
PROGRAM PARTICIPATION

SOME ASPECTS OF SIPP

No. 09

R. A. Herriot and D. Kasprzyk
Bureau of the Census

U. S. Department of Commerce BUREAU OF THE CENSUS

Survey of Income and Program Participation

SOME ASPECTS OF THE SURVEY OF INCOME
AND PROGRAM PARTICIPATION

No. 8601 - 9

April 1986

ACKNOWLEDGEMENT

This paper was compiled and edited by Daniel Kasprzyk and Roger A. Herriot for presentation to the Census Advisory Committees of the American Statistical Association and on Population Statistics at the Joint Advisory Committee Meeting, April 10-11, 1986. Because of the addition of more current information, this version of the paper differs in minor ways with the paper distributed to the Advisory Committees.

The paper represents the collective efforts of many Census Staff whose areas of responsibilities include various aspects of the Survey of Income and Program Participation. In particular, the following individuals contributed to this review of SIPP; John Coder, John McNeil, Enrique Lamas, Paul Ryscavage, Donald Hernandez, David McMillen, Edith McArthur, Kathleen Short, and Delma Frankel, all of the Population Division, Raj Singh from Statistical Methods Division, Chester Bowie, Barry Fink, and Donna Riccini, all from the Demographic Surveys Division, and John Spendlove from Field Division. A special thanks also goes to Hazel Beaton for her patience and accurate typing of the paper under a very tight schedule.

TABLE OF CONTENTS

		Page
I.	Background	1
II.	Planned Data Collection Metholodogy	2
	Sample Reduction	2
	Content Determination	4
	1986 Panel Topical Modules	4
	Interview Mode	5
	SIPP Telephone Interview Pretest	5
	SIPP National Telephone Interview Test	6
	Experiments with Dependent Interviewing	6
	Wave 7 Asset and Liability Feedback Study	6
	Industry and Occupation Coding in the 1986 Panel	7
	Signed Release Feasibility Study	8
III.	SIPP Data Collection Process	9
	The Interviewer	9
	The Respondent	12
	Data Processing	14
IV.	Current and Past Income Measurement Evaluation	14
	A Review of Nonsampling Error Issues in the SIPP	14
	Unit and Person Noninterview	15
	Cross-Sectional Item Nonresponse	21
	Longitudinal Item Nonresponse	21
	Evaluation of Survey Reports of Income	24
	Evaluation of Labor Force Estimates	26
	SIPP Gross Flow Data	27
	Reinterview Program	31
V.	Planned Evaluation of SIPP Data	31
	Center for Survey Methods Research: Multi-State Multi-Program Record Check Study	32
	ASA/NSF/Census Bureau Fellowship Program	32
	Federal Users of SIPP Data	33
VI.	Analysis and Products at the Census Bureau	34
	Reports and Files	34
	Longitudinal Research File	35
	Longitudinal Concepts	37
	Longitudinal Processing	38

VII. Activities Related to User Needs..... 38

VIII. Other Topics..... 40

Statistical Estimation for Longitudinal Concepts..... 40
Composite Estimation..... 40
Variance Estimation..... 41
Cross-sectional Estimation..... 41
SIPP Wealth Data..... 41
Matching Activities..... 41
Other Analysis (Not covered elsewhere)..... 42
Variance Reduction through Post-stratification..... 43

References

PREFACE

This paper provides background information and references to work in progress and work completed in a number of areas important to the Survey of Income and Program Participation (SIPP). The paper briefly describes topics related to data collection; indicates briefly what we know about income measurement in the SIPP; identifies several planned evaluation activities; discusses analysis and data products; describes past and ongoing activities related to user needs; and discusses a number of areas in which research has been initiated.

TABLES

Table 1 Sample Size Summary for the 1984, 1985, and 1986 SIPP Panels 3

Table 2 Workload and Turnover Comparisons by Program by Regional Office 10

Table 3 Household Interview Time by Size of Household and Percent of Self-Response for Self-Interviews: 1984 SIPP 13

Table 4 SIPP Noninterview Rates and Sample Loss 16

Table 5 New Noninterview Rates by Wave: 1984 Panel 16

Table 6 Response Patterns of SIPP Original Sample Persons for the First Five Interviews of the 1984 SIPP Panel..... 20

Table 7 Item Nonresponse Rates for SIPP and March 1985 CPS, for Selected Income Types 22

Table 8 Longitudinal Item Nonresponse Rates for Amounts of Selected Income Types: 1984 SIPP Panel 12-month Summary 23

Table 9 Comparisons of Estimated Numbers of Income Recipients and Estimated Aggregate Income Amounts Received for Selected Income Types: SIPP vs. Independently Derived Estimates vs. the Current Population Survey 25

Table 10A Inconsistencies in 1984 Panel Demographic Data: Wave-to-Wave Race Change..... 36

Table 10B Inconsistencies in 1984 Panel Demographic Data: Wave-to-Wave Sex Change..... 36

Table 10C Inconsistencies in 1984 Panel Demographic Data: Wave-to-Wave Age Change..... 36

FIGURES

Figure 1 New Type A's and D's for SIPP Waves 1-7 17

Figure 2 Comparing SIPP and PSID Sample Loss 18

Figure 3 SIPP-CPS Labor Force Comparisons, I 28

Figure 4 SIPP-CPS Labor Force Comparisons, II 29

Figure 5 SIPP-CPS Labor Force Comparisons, III 30

ATTACHMENTS (Available Upon Request)

Attachment 1	The Goals and Objectives of the Survey of Income and Program Participation
Attachment 2	An Overview of the Survey of Income and Program Participation: Update 1 SIPP Working Paper No. 8401
Attachment 3	Office of Management and Budget SIPP Advisory Committee
Attachment 4	Social Science Research Council Committee on the SIPP
Attachment 5	SIPP Product Plans
Attachment 6	Available SIPP Reports, Working Papers, and Compilations

I. BACKGROUND

In October 1983, the Bureau of the Census conducted the first interviews in the Survey of Income and Program Participation (SIPP). The SIPP is a nationwide survey designed to provide accurate, comprehensive information about those factors affecting the economic situation of persons and households which are relevant for government policy decisions. A discussion of this goal and the objectives which must be achieved to meet this goal can be found in attachment 1.

SIPP was designed to improve reporting of income and program-related data and do it in a way that would allow the analysis of changes over time at the microlevel. The design also had to accommodate the collection of a large quantity of information in a flexible manner that allowed some information to be collected more frequently than other information. These requirements were met principally by using a survey design in which the same people are interviewed more than once. Persons at households selected for a sample panel are interviewed about their income and other topics once every 4-months for approximately 2 1/2 years. Sample persons are interviewed at new addresses if they move. Also any other persons that they move in with, or vice versa, are interviewed. In this way a highly detailed record is built up over time for each person and household in a sample panel. This design minimizes the need for sample persons to recall most of the information for longer than a few months and reduces the number of questions asked in one interview. To further enhance the estimates of change, particularly year-to-year change, a new sample panel is introduced every year; consequently, two and sometimes three panels are in the field concurrently. Thus, SIPP is a rotating panel survey, with a new panel begun each year that is interviewed at 4-month intervals for 2 1/2 years. Attachment 2, "An Overview of the Survey of Income and Program Participation: Update 1," provides a general introduction to the SIPP, featuring descriptions of the design features, survey content, and operational procedures.

This paper provides a review of several aspects of the SIPP:

1) Planned data collection methodology;

2) The data collection process;

3) Current and past income measurement evaluation;

4) Planned evaluation of SIPP data;

5) Analysis and products at the Census Bureau;

6) Activities related to user needs; and

7) A potpourri of additional topics.

II. PLANNED DATA COLLECTION METHODOLOGY

The continuous nature of data collection in SIPP results in a large part of the budget being assigned to field costs. Since interviews are taken every 4-months, new questionnaire modules must be regularly developed by the Census Bureau. In addition, continuous data collection offers the possibility of experimenting with different collection methodologies. This section reviews the SIPP sample size and describes the sample reductions taken to balance the SIPP budget. A description of the process to determine content for the SIPP and decisions made concerning the 1986 Panel topical modules will also be provided. Finally, experiments in data collection methodologies included in the SIPP Panels will be described.

Sample Reduction

In FY 1985, the final base budget allocation for SIPP was $1.2 million less than the amount of funding originally requested to conduct the survey. This budget reduction required long-term cost avoidance measures which were instituted by:

1. Introducing a sample reduction beginning in March 1985 for the 1984 Panel by deleting 850 interviewed households from each of the four rotation groups (4 x 850 = 3,400 interviewed households). This sample reduction yielded a 17.8 percent cut in sample size. Both self-representing (SR) PSUs and nonself-representing (NSR) PSUs were subject to the cut.

2. Introducing a sample reduction beginning in February 1985 for the 1985 Panel by deleting 1,320 interviewed households from each of the four rotation groups (4 x 1,320 = 5,280 interviewed households). This sample reduction was implemented by dropping 3,120 interviewed households from the SR PSUs and dropping 54 NSR PSUs (2,160 interviewed households).

In FY 1986, the Gramm-Rudman-Hollings Act necessitated additional long-term cost avoidance measures which were instituted by:

1. Introducing an additional sample reduction beginning in February 1986 for the 1985 Panel by deleting 490 interviewed households from each of the four rotation groups (4 x 490 = 1,960 interviewed households). This sample reduction yielded a 15 percent cut in sample size. Both SR and NSR PSUs were subject to the cut.

2. Introducing an identical sample reduction as mentioned above for the 1986 Panel.

3. Accelerating interviewing on the 1984 Panel so that the second annual roundup interviews (Wave 9) will end a month early. This acceleration will shorten the longitudinal duration of the 1984 Panel by 1 month, and result in the Wave 8 topical module questionnaire being administered to three-fourths of the sample.

Table 1 provides a summary of the sample size for the 1984, 1985, and 1986 SIPP Panels.

TABLE 1. -- Sample Size Summary for the 1984, 1985, and 1986 SIPP Panels

SUMMARY

1984 Panel (Started October 1983)

Eligible households initially assigned for interview	20,897
Interviewed households remaining after seven waves of interviews (after sample reduction for budget reasons)	14,902
Sample loss due to household nonresponse after seven waves of interviews	21.0%

1985 Panel (Started February 1985)

Eligible households initially assigned for interview	14,306
Interviewed households remaining after three waves of interviews	13,077
Sample loss due to household nonresponse after three waves of interviews	13.0%

1986 Panel (Started February 1986)

Eligible households initially assigned for interview	12,100

Content Determination

A SIPP Advisory Committee, chaired by a representative from the Office of Management and Budget (OMB), advises and recommends to the Census Bureau changes in SIPP content, particularly as these changes relate to data required for policy analysis. The OMB Advisory Committee includes representatives from over twenty Federal agencies. After advising and assisting on the development of the SIPP core questionnaire on income and program participation, the committee has served principally as an advisory board for, first, identifying topics for the topical modules for each panel, and then identifying specific questions within each topic. Typically, a Census Bureau representative chairs a subcommittee which is formed to develop each topical module questionnaire. Topics are considered and debated within the subcommittee; after determining the topics to be covered, a questionnaire is developed. The draft questionnaire, after receiving subcommittee approval, is brought before the full Advisory Committee for further discussion. At the end of this process all Federal agencies represented on the committee have had an opportunity to discuss their requirements for data among their peers, and participate in the development of the questionnaire. A pretest of the draft questionnaire on a small sample (usually around 100 households) follows. The results of the pretest produces a revised questionnaire which again is discussed with subcommittee members. This process of content determination and questionnaire design in SIPP, followed for each SIPP Panel thus far, is somewhat unwieldy, but it does allow ample opportunity for discussion and criticism. Members of the SIPP Advisory Committee are listed in attachment 3.

In the early stages of the development of SIPP, Census Bureau staff recognized the potential of SIPP to the academic community and have solicited comments from this community on the design, methodological problems, and content of the SIPP. The vehicle for this solicitation is the Social Science Research Council's (SSRC) Committee on the SIPP. This Committee, sponsored by a grant from the National Science Foundation, has developed workshops and symposia, and encouraged analysts to reflect on whether the SIPP data meet their needs. The committee has been instrumental in expanding the academic community's knowledge of SIPP and has helped in clarifying the content issues for SIPP. Attachment 4 provides a listing of the members of the SSRC SIPP Committee.

1986 Panel Topical Modules

Census Bureau plans for the content of the 1986 Panel evolved from these requirements:

1. The complexity of the interviewing task, both for the interviewer as well as respondent, should be reduced;

2. Deficiencies in measures required for longitudinal analysis ('left censoring') should be corrected;

3. Stability of the questionnaire design should be maintained to minimize data processing complications.

As a result, the SIPP OMB Advisory Committee assisted the Bureau staff in developing a design for the 1986 Panel by reviewing the content of the core and topical modules. They proposed 1) that core questions for the 1986 Panel remain virtually unchanged and 2) that questions on personal history be organized in a topical module as close to the first interview as possible--this baseline interview will take place in Wave 2 of the 1986 Panel. This module contains bounding information on the duration of status at the time of recruitment into the panel for several areas--retrospective employment, fertility, program participation history, marital history, education and training history, and kinship. The committee also proposed that questions on assets and liabilities be organized as modules to be asked in Waves 4 and 7, and that the tax information and reconciliation of annual income and within year income reports be asked in Waves 5 and 8.

Thus, Waves 3 and 6 were identified as interview waves available for varying content. In Wave 3 the OMB Advisory Committee recommended the following topics: 1) child care arrangements; 2) child support agreements; 3) support for nonhousehold members; 4) job offers; 5) health status and utilization of health care services; 6) long term care; and 7) disability status of children.

While in Wave 6, the Committee recommended: 1) child care arrangements; 2) child support agreements; 3) support for nonhousehold members; 4) work-related expenses; 5) housing costs; and 6) energy usage.

Appendix B in attachment 2 provides a summary of the content in all SIPP Panels.

Interview Mode

The SIPP has followed the practice of many surveys and conducted most interviews (approximately 95 percent) in person. This procedure, a legacy of the Income Survey Development Program (ISDP), the predecessor to SIPP, is more expensive than telephone or mail surveys, but usually results in more complete data. Since the primary goal of the SIPP was to improve the reporting of income and program participation in Federal programs, alternatives to personal interviews were never seriously considered. Because of the rising costs of a personal interview, and an uncertain budget, it is apparent that the Census Bureau must consider the possibility of conducting SIPP interviews by telephone. A telephone interview pretest was conducted in June 1985 to provide the necessary experience to develop a national telephone interview test to be conducted in the late summer to early fall of 1986.

SIPP Telephone Interview Pretest--

A SIPP Telephone Interview Pretest was conducted in June 1985 to assess the feasibility of conducting SIPP interviews by telephone. The sample for the pretest was selected from cases in the Atlanta and Chicago regional offices that were dropped as a result of the 1984 panel sample reduction and identified as respondents who would accept a telephone interview by the interviewers. The pretest was conducted through the Atlanta and

Chicago regional offices by experienced interviewers. Some interviews were conducted at the regional offices with observers monitoring the interviews, but the majority of the interviews were conducted at the interviewers' home.

The observers of the test concluded that it was feasible to conduct SIPP interviews by telephone with only minor changes in the questionnaire structure. The analysis of the item nonresponse rates for the pretest is still in progress.

SIPP National Telephone Interview Test--

A SIPP National Telephone Test will be conducted in August to November 1986. The test is designed to provide better estimates of data quality and cost per case for telephone interviewing versus personal visit interviewing.

The 1986 test will be conducted during the last two rotations of Wave 2 and the first two rotations of Wave 3 of the 1986 Panel. Households within 50 percent of the segments in the targeted rotations will be designated as maximum telephone interview cases. The remaining 50 percent will be maximum personal visit cases. Interviewers will conduct almost all of the telephone interviews from their homes.

The analysis of the data from the test will focus on comparing the telephone interview data with the personal visit interview data. Specifically, estimates of various characteristics, item nonresponse rates, and the cost data for the two interview modes will be compared.

Experiments with Dependent Interviewing

Since interviews are conducted with the same persons at short time intervals, the SIPP can feed back information to respondents which had been previously collected. This has the obvious benefit of reducing response error in the measurement of change. As SIPP provides two measurements of net worth during a panel and multiple measurements of industry and occupation, the use of dependent interviewing to mitigate the effect at the microdata level of two independent interviews is worthy of study. Two studies are described below.

Wave 7 Asset and Liability Feedback Study--

Household surveys represent a major data source to study the composition and distribution of household wealth. SIPP provides a recurring series on household wealth. Data on assets and liabilities are important in determining program eligibility and assessing the economic situation of families. SIPP is scheduled to collect asset and liability information two times for each panel at one year intervals. Detailed and comprehensive questions concerning the ownership and amounts of assets and liabilities were included in the Wave 4 topical module of the 1984 Panel (collected in September through December 1984). These items were updated one year later in Wave 7.

Viewed longitudinally, collecting asset and liability data two times per panel provides the only available survey microdata on consumer savings, i.e., the change in asset equity. Response errors in the reporting of asset and liability amounts are expected to be present and affect micro-level measures of change, or savings. In order to obtain more consistent measures of consumer savings, research to feed back information to respondents during an interview was proposed. Specifically, information on asset and liability values collected in Wave 4 was provided to one-half of the respondents interviewed in the Wave 7 interview.

The rationale for the feedback proposal is that respondents will provide more accurate estimates of change if they are first reminded of the amount they reported holding at the beginning of the period. If respondents have knowledge of the amount of the change in asset value and are reminded of their beginning balance, then their reporting of their current balance may be consistent with the true amount of change over the period.

To test the feedback approach, a split sample was implemented; approximately one-half of the eligible Wave 7 households were given the feedback form, while the remaining sample households were asked the asset and liability items without the previously reported information. This provided a control group to compare the feedback information.

No microlevel administrative record sources are available to benchmark savings estimates from SIPP. As a result, no definite conclusion can be reached on the accuracy of the data. There are, however, several ways to judge the reasonableness of the data collected and to draw inferences about the quality of the feedback-based data. For both samples, savings estimates at the person and household level can be compared to the 1985 Panel estimates to see if any systematic differences arise between the samples at the aggregate level.

In addition, savings behavior is related to other economic factors. Savings are expected to be related to employment patterns, income level of person and household, and recipiency of certain income types. Information in SIPP on the person and household can be used to assess the reasonableness of the data. For example, income is positively correlated with savings, while periods of unemployment are expected to be negatively correlated with savings. Other factors such as coverage by pension plan, life insurance policies, and health insurance coverage affect savings. Comparing savings patterns of individuals of each split sample using other economic information may give an indication of the impact of the feedback procedure. The results of this research will indicate whether it is feasilible to measure consumer savings using SIPP and whether an update of the asset and liability topical module is necessary for future panels.

Industry and Occupation Coding in the 1986 Panel--

During the 1984 and 1985 SIPP panels the industry and occupation data were collected independently during each interview even though the individual had not changed employers. This procedure acknowledges the fact that an

employee's duties may change from time to time and allows these changes to be recorded. Sufficient change in duties could result in a change in the person's occupation classification from interview to interview even though the employer has not changed.

The independent collection of industry and occupation data has, however, several problems. Undue variation in occupation classification can result when respondent descriptions of duties vary slightly or when the interpretation of the written description varies between the clerical staff members assigning the classification codes based.

Research into this problem has provided some estimates of the number of times occupation and industry classifications change from interview to interview for persons with the same employer. About 40 percent of these persons change 3-digit occupation codes between interviews and 20 percent change 3-digit industry codes. About 50 percent of these changes were termed major, i.e., the person changed one of 10 major occupation groups or one of 14 major industry groups.

A change in data collection procedures for occupation and industry was implemented in the 1986 Panel. This modification was designed to reduce changes in occupation and industry codes resulting from random response error and clerical interpretation, and to reduce interview time. The modification introduces a "screener" question that asks if activities or duties have changed during the past 8 months. A negative response eliminates the detailed occupation and industry questions. The occupation and industry classifications would then be derived from responses given in the initial interview.

It is important to note that while this change is being made for the 1986 Panel, industry and occupation data from the 1985 Panel collected during the same time period as the 1986 Panel will be collected independently each wave, giving rise to a natural experiment embedded in the two panels.

Signed Release Feasibility Study

Employer-provided benefits have become an important part of employee compensation. To obtain estimates of the value of these benefits, information on the coverage and employer contributions to health insurance, pension plans, and life insurance plans are required. A feasibility study of the collection of data from employers will be conducted by the Census Bureau. In this study, approximately 300 households in the last rotation of the 1984 SIPP Panel will be used in a pretest of the procedure. A signed release will be obtained from respondents at the interview and a form will be sent to their employers. The purpose of the test is to determine the response rate and to observe the problems encountered in obtaining the data. If an adequate response is obtained, the Census Bureau will proceed in developing this project for a future panel of SIPP. This study has been delayed for approximately one year.

III. SIPP DATA COLLECTION PROCESS

The SIPP questionnaire looks very large and complex; it must be in order to identify and probe for both major and minor sources of income. Associated with the questionnaire is the perceived burden on both the interviewer and the respondent. This section provides information concerning the interviewer, respondent, and the processing system.

The Interviewer

Based on at least 30 years of research at the Census Bureau as well as other survey organizations, we know a relationship exists between interviewers and the data they collect. The importance of this relationship varies by the amount of interaction between interviewer and respondent. SIPP surely must be, in terms of Census Bureau collection, the survey in which interviewer respondent interaction is maximized--the relationship is both intense and long-term. The success of the survey no doubt depends on the quality of that relationship. Thus, interviewer turnover rates are especially important for the SIPP. Table 2 provides workload and turnover comparisons by census program and regional office. Note that despite the perceived burden on the interviewer of a long and complicated form, SIPP interviewer turnover rates are lower than the Current Population Survey and the National Crime Survey turnover rates in six out of the twelve regional offices.

TABLE 2. -- Workload and Turnover Comparisons by Program by Regional Office

Regional Office	Average Monthly Workload				Interviewer Turnover Rates			
	CPS	NCS	CE	SIPP	CPS	NCS	CE	SIPP
National	73,538	14,088	3,752	7,775	7.0	8.9	5.6	10.3
Boston	9,299	1,096	344	567	4.5	7.0	10.5	4.2
New York	5,017	858	184	514	15.0	11.5	0.0	32.7
Philadelphia	6,599	1,165	299	736	5.1	0.0	0.0	1.7
Detroit	5,863	1,347	289	552	3.4	27.5	0.0	0.0
Chicago	3,985	789	198	565	5.3	5.6	0.0	4.8
Kansas City	4,527	1,066	407	664	2.0	2.2	0.0	12.5
Seattle	5,477	1,085	304	656	9.1	7.5	12.1	5.8
Charlotte	6,840	1,338	285	628	10.6	7.5	5.9	11.4
Atlanta	6,053	1,601	409	854	7.2	12.2	13.8	10.2
Dallas	7,092	1,735	377	881	8.0	1.8	5.6	14.6
Denver	8,304	1,085	340	565	4.5	25.0	5.1	14.6
Los Angeles	4,482	923	316	593	10.2	6.7	9.8	8.8

Workload is expressed as the average number of households assigned monthly during the period July - October 1985.

Turnover is the percent of the interviewer staff who resigned during the period October 1985 - January 1986.

Source: Methods Research Branch
Field Division
Bureau of the Census

Because the survey depends on the quality of the interview-respondent relationship, a project has begun to study interviewer characteristics as they may affect response rates in the SIPP. As there was no source for this information from personnel or field office records, SIPP supervisors in each of the regional offices were asked to fill in a short questionnaire for each of the interviewers in their offices who worked or had worked on SIPP since it began in October 1983.

The information on these questionnaires was keyed and an SPSS-formatted file has been created containing a record for each interviewer for whom a questionnaire was filled. From this a profile of SIPP interviewers will be created. The information on interviewers will also be merged with a file of all SIPP respondents--so associated with each SIPP respondent on the data file will be the characteristics of the interviewer who conducted each interview.

Some of the things which will be described include: whether there is a relationship between a change in interview from one wave to the next and response rates in terms of attrition as well as in terms of specific data collected; what impact size of the interviewer's workload has upon their type A and D rates; the effect of number of years of experience as an interviewer in total and specifically on SIPP upon an interviewer's success as measured by response rates; and if interviewers work on other surveys in addition to the SIPP, what kind of impact does that have upon their performance--that is in terms of priorities on work completion and in terms of conflicting instructions which might lead to confusion.

Although there is a long history of concern and research on interviewers and training and how they relate to the data that are collected, no recent report on interviewer characteristics is available at the Bureau. We intend to provide some descriptive information on this topic.

Interviewer training is a critical component of the data collection process. Without adequate training which explains forms, procedures, and reasons for collecting the data, the interviewers quite simply cannot do their jobs. Although the general feeling about SIPP training materials and procedures was that they were more than adequate, an independent review and evaluation by an expert consultant was initiated. The obvious goal of the review was to obtain recommendations for improvements in SIPP interviewer training and materials.

A preliminary report based upon a review of the interviewer's manual, self-study materials, and training guides has been received (Holt, 1985). Recognizing that the interviewers were the ultimate users of these materials, the consultant prepared this preliminary report using her perception of their needs. The recommendations included more use of the interviewer's manual during training so that interviewers would be more familiar with--and more likely to refer to--the manual when necessary. She also recommended the use of more specific examples and exercises for particular situations in the training materials.

The final report for this project covers the consultant's evaluation and recommendations for improvements after observing several training sessions and after going into the field with an interviewer to observe an actual interview. Copies of a draft of this report are being circulated for comments (Holt, 1986).

The Respondent

The burden of the respondent can be roughly looked at in two ways. First, one can calculate the average number of responses to questions in the questionnaire. The number of responses is largely a function of the economic experiences of the respondent and, of course, can vary widely depending on the complexity of his/her economic situation. In Wave 2 of the 1984 Panel, an average of 70 responses per questionnaire occurred; in Waves 6, 7, and 8 of the 1984 Panel, an average of 116 responses, 103 responses, and 110 responses per questionnaire occurred. The difference between Wave 2 and the other waves can be attributed to the fact that Wave 2 did not include a topical module. These counts do not include any information collected on the control card, nonsource coded information collected on the cover page of the questionnaire, or employer/business identifying information collected in the Earnings and Employment sections of the questionnaire.

The second way of looking at burden on the respondent is examining the length of SIPP interviews and the respondent type. The time required to conduct a SIPP interview is outlined in table 3 for interviews 1, 2, and 5. The first two interviews did not contain a topical module while the fifth included an extensive group of questions covering child care, welfare history, reservation wage, etc. The fifth interview included all of the basic "core" questions common to the second interview as well as the topical module. The interview times for the first interview reflect the additional time required for the initial contact and the "newness" of the questionnaire and survey procedures. The lower times in the second interview reflect the experience obtained by the interviewers in the first interview and the design of the questionnaire which updated the income profile developed in the previous visit.

The last row of the table shows that the proportion of person interviews conducted with the individual for whom the questions applied (self-response) remained about the same through the first five interview periods.

TABLE 3. -- Household Interview Time by Size of Household and Percent of Self-Response for Self-Interviews: 1984 SIPP

(Time in Minutes)

SIZE OF HOUSEHOLD 1/	INTERVIEW NUMBER		
	1	2	3 2/
Total	43	28	31
One	29	18	18
Two	44	29	32
Three	57	39	42
Four	70	49	53
Five	83	58	61
Six	98	70	69
Seven or more	113	78	100
Percent self response . .	62	62	61

1/ Persons 15 years old and over.

2/ Includes time for additional questions in the topical module.

Data Processing

The keying of the SIPP data is currently being done in each of the 12 Census Bureau Regional Offices. The data are transmitted, on a flow basis, from each office into a main frame computer located at the Census Bureau headquarters. An extensive data acceptance edit is run on each transmission and rejects are transmitted back to the originating office. This procedure puts the correction of rejects and the responsibility for missing data in the same office that is responsible for managing data collection. On the average, a regional office sends eight original data transmissions per panel each month and must recycle reject corrections for each transmission four times.

At time of initial interview the only control information available is the address of each household selected for the sample. Some reporting, interview or noninterview, for each of these households is required. After the first interview is complete, a person-level control file is created for the second interview. Every person interviewed in the first wave must be accounted for in the second. New sample persons identified in the second interview will become part of the person-level control file for the third interview and so on. Each person on the file is uniquely identified by a five-part control number made up of the PSU Code, Segment Code, Serial Number, Entry Address ID and Person Number. This control number on the data file is matched to the corresponding number on the control file and the basic demographic characteristics (age, sex, race) listed on the two files are compared to verify that the match is correct.

No interview wave is closed out until every person listed on the control file is accounted for in the current wave of interviewing. This strict person-level control is somewhat difficult and expensive to maintain but preliminary work with matching files from mutliple waves indicate that the quality of our longitudinal data is greatly improved over our experience with the Income Survey Development Program (ISDP). See the section "Longitudinal Research File" for more information.

IV. CURRENT AND PAST INCOME MEASUREMENT EVALUATION

Income measurement evaluation is related to understanding the effects of nonsampling error on SIPP estimates. At this time we have no experimental evidence showing the magnitude and effects of nonsampling error in the SIPP. We have, however, begun to develop a better understanding of the issue along several dimensions.

A Review of Nonsampling Error Issues in the SIPP

A paper prepared for presentation at the Census Bureau's Annual Research Conference reviews some of the SIPP design decisions affecting the quality of the various forms of estimates produced from SIPP data (Kalton, McMillen, and Kasprzyk, 1986). This paper reviews two major issues of nonsampling errors in panel surveys--panel attrition and panel conditioning. It then

examines nonsampling errors that arise from major design decisions such as 1) the decision to use a panel design; 2) the length of reference period; 3) respondent rules; 4) following rules; and 5) the mode of data collection.

Finally, the paper discusses how various design decisions affect the measurement of change. When available, SIPP data are used to illustrate points. Additional work on this topic will be available by September 1986, as a paper is planned for presentation at the 1986 Meetings of the American Statistical Association (McMillen, Kasprzyk, and Kalton, 1986).

Unit and Person Noninterview

Unit and person nonresponse is one aspect of nonsampling error in which data are readily available from field operations. The effect of nonresponse on survey estimates, however, is an area where little data are available. As will be described below, some information on the distribution of persons who leave the SIPP sample has recently become available.

In SIPP, nonresponse can be measured in several ways (Bailey, Chapman, and Kasprzyk, 1985). The routine way is to consider the total number of eligible households assigned including the type A noninterviews (household noninterviews including refusals, no one at home, etc.) for Wave 1 as denominator. Of course, the numerator is the total number of type A's in the survey.

In SIPP, an additional form of unit noninterview exists because survey procedures call for following all people who lived at the sample address at the time of the first interview. Thus, a type D noninterview household is defined as a household of one or more original sample persons who can not be followed to their new address or who moved beyond 100 miles from a SIPP PSU. Table 4 provides the cumulative type A and type D rates by wave for the 1984 and 1985 Panels.

For the 1984 and 1985 Panels many cases which are a noninterview in one wave are reassigned in the next wave. Reassignment is discretionary and depends in large part on field staff's subjective assessment of the probability of conversion to interview status. Thus, there are households in each wave eligible for interview which have been noninterviews in previous waves. This gives rise to another way of looking at "cross-sectional" noninterview rates--that is, one can look at the new noninterview rate by wave of the type A and type D households. Table 5 and figure 1 present this information for the 1984 Panel; as the length of time the sample is in the field increases, the new noninterview rate decreases.

The Census Bureau prides itself on its ability to obtain high response rates in its current surveys programs. The SIPP, by nature of its design, should expect its cumulative nonresponse rate to increase after each interview. It is difficult to imagine how the SIPP cumulative nonresponse rate should remain the same or decrease at the conclusion of each wave of interviewing. To put the SIPP nonresponse experience in perspective,

figure 2 compares the SIPP and Panel Study of Income Dynamics (PSID) sample loss after seven interviews. SIPP starts with a much lower noninterview rate than the PSID; after seven interviews, the rate of sample loss for both surveys is, however, approximately the same. It should be noted that PSID-SIPP comparisons are for the same number of interviews over a substantially different period of time--the SIPP covering 7 interviews over approximately two years and the PSID covering 7 interviews over seven years. Additional discussion of noninterview rates, comparisons to other panel surveys, and a discussion of the methods Census staff are using to maintain respondent cooperation will soon be available in a SIPP Working Paper (Nelson, McMillen, and Bowie, 1986).

TABLE 4. -- SIPP Noninterview Rates and Sample Loss

	1984 Panel			1985 Panel		
Wave	Type A Rate	Type D Rate	Sample Loss	Type A Rate	Type D Rate	Sample Loss
1	4.9%	----	4.9%	6.7%	----	6.7%
2	8.3%	1.0%	9.4%	8.5%	2.1%	10.8%
3	10.2%	1.9%	12.3%	10.0%	2.8%	13.0%
4	12.1%	2.9%	15.4%			
5	13.4%	3.5%	17.4%			
6	14.9%	4.1%	19.4%			
7	15.6%	4.9%	21.0%			

Type A noninterviews consist of households occupied by persons eligible for interview and for whom a questionnaire would have been filled if an interview had been obtained. Reasons for type A noninterviews include: no one at home in spite of repeated visits, temporarily absent during the entire interview period, refusal, and unable to locate a sample unit.

Type D noninterviews consist of households of original sample persons who are living at an unknown new address or at an address located more than 100 miles from a SIPP PSU, provided a telephone interview is not conducted.

TABLE 5. -- New Noninterview Rates by Wave: 1984 PANEL

Wave	New Type A Rate	New Type D Rate
2	3.5%	1.0%
3	3.3%	1.2%
4	2.7%	1.0%
5	1.9%	0.7%
6	1.8%	0.5%
7	1.1%	0.9%

Figure 1

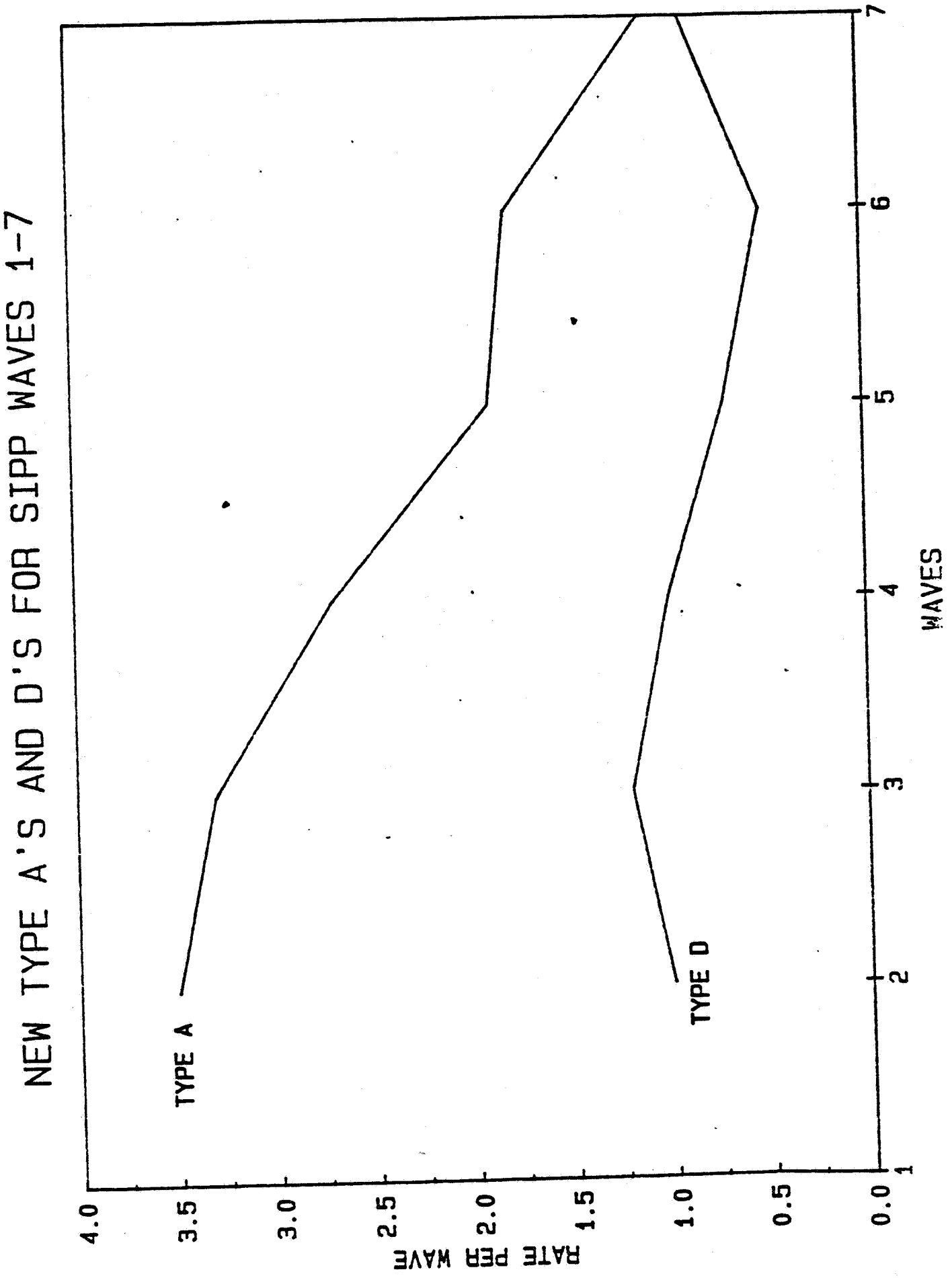
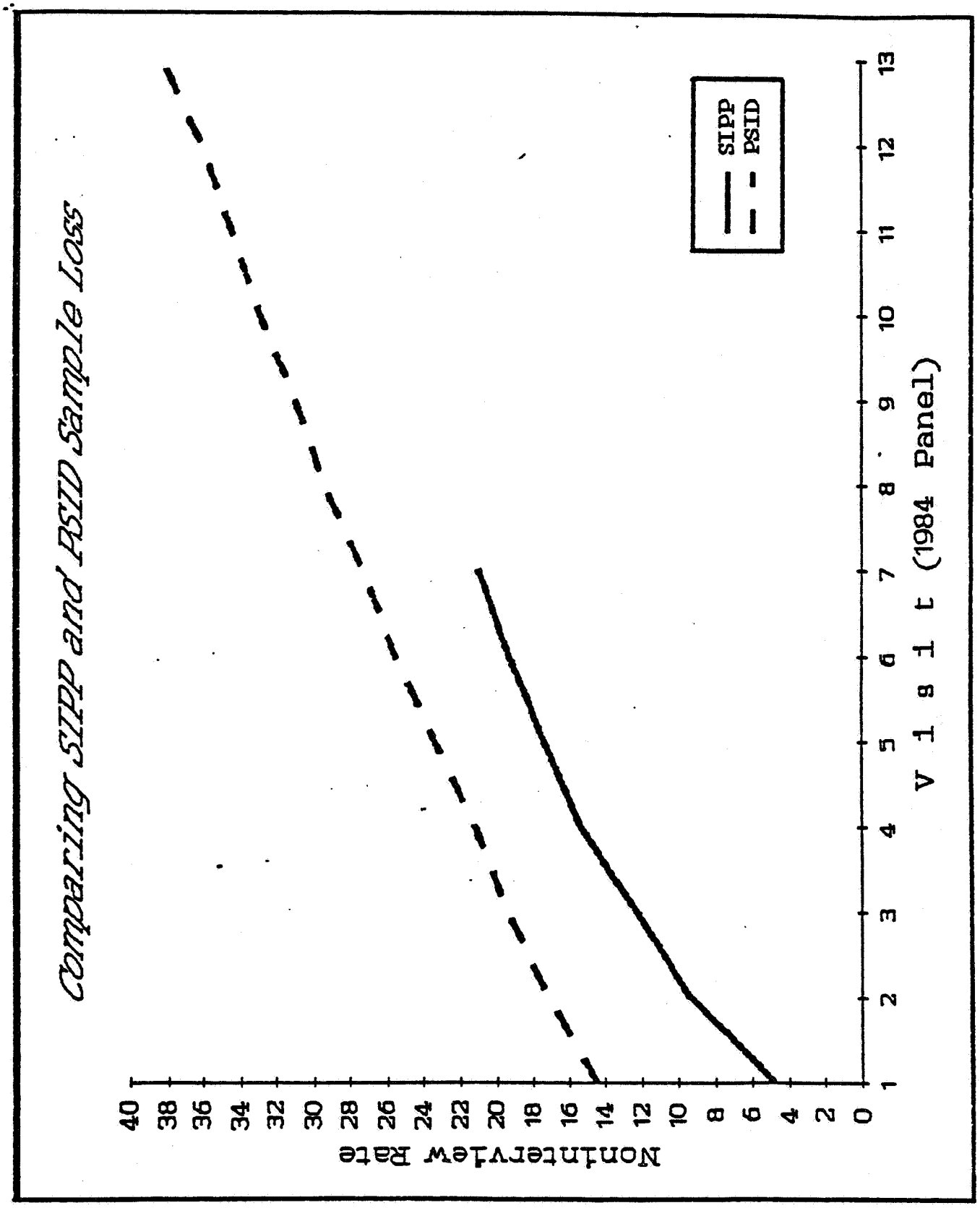

NEW TYPE A'S AND D'S FOR SIPP WAVES 1-7
TYPE A
TYPE D
RATE PER WAVE
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
1
2
3
4
5
6
7
WAVES

Figure 2


Comparing SIPP and PSID Sample Loss
Nonintervlew Rate
40
38
36
34
32
30
28
26
24
22
20
18
16
14
12
10
8
6
4
2
0
1
2
3
4
5
6
7
8
9
10
11
12
13
V i s i t (1984 Panel)
SIPP
PSID

Another way of viewing response rates in the SIPP is to look at them on a "person" basis as opposed to a household basis; that is, consider sample loss in terms of the reduction in the numbers of initially interviewed sample persons over the time those persons were eligible for interview. McArthur and Short (1985) have described the SIPP sample loss for three interviews; recently they have included two additional waves of data in their research. Table 6 from Kalton, McMillen, and Kasprzyk (1986) provides a summary of response patterns observed in SIPP after five interviews; additional tables which present reasons given for the non-interviews and characteristics of persons who leave the sample are available (McArthur and Short, 1986).

Compensating for nonresponse described above can take place through either weighting adjustments or imputation. The choice is not obvious and is complicated by the substantial amount of information available from other waves of a panel survey. The complications occur because it is not clear how best to use the additional data and then, because different analyses require different subsets of the data, what is the appropriate number of sets of weights to make available to analysts. The Survey Research Center (U. of Michigan) has begun to study this topic (Kalton, 1985; Kalton, Lepkowski, and Lin, 1985). Additional work is in progress and results will be reported in August 1986 at the meeting of the American Statistical Assocaition (Kalton and Miller, 1986). Preliminary results, taking into account the longitudinal relationship between the variable and simple prediction models, show imputation can be more efficient then weighting. However, the practical realities of developing good imputation models for a missing wave is a significant undertaking, and furthermore, the use of different sets of weights for different analyses can lead to inconsistent results. Research will continue this year but the likely results may very well be a combination of the two options.

TABLE 6. -- Response Patterns of SIPP Orginial Sample Persons For the First Five Interviews of the 1984 SIPP Panel 1/

		Number	Percent
1. Response every interview (5 interviews)			
Pattern:	XXXXX	19878	79.08
2. Apparent attrition cases		3459	13.76
Patterns:	XXXX0	964	3.83
	XXX00	768	3.06
	XX000	811	3.23
	X0000	916	3.64
3. First and fifth interviews but one intervening interview missing		863	3.43
Patterns:	XXX0X	413	1.64
	X0XXX	148	0.59
	XX0XX	302	1.20
4. First and fifth interviews, two or more intervening interviews missing		165	0.66
Patterns:	X000X	30	0.12
	X0X0X	18	0.07
	XX00X	75	0.30
	X00XX	42	0.17
5. Fifth interview missing and one or more intervening interviews missing		196	0.78
Patterns:	X0XX0	29	0.12
	X0X00	61	0.24
	X00X0	22	0.09
	XX0X0	84	0.33
6. Left the universe (deceased, institutionalized, living in armed forces barracks, moved overseas)		577	2.29

1/ The universe for the table consists of all persons in rotation groups 1, 2, and 3 who were 15 years or over at the time of the first interview and for whom a personal interview was conducted (either self or proxy interviews) during the first wave of the 1984 SIPP Panel, and who were designated for interview for all five interviews. The symbol "X" represents a successful interview and the symbol "0" represents no interview (either no household interview or no personal interview).

Cross-sectional Item Nonresponse

Discussions of the levels of item nonresponse in the SIPP have regularly occurred at the meetings of the American Statistical Assocaition (Coder and Feldman, 1984; Lamas and McNeil, 1984; McMillen and Kasprzyk, 1985). These reports have focussed on cross-sectional item nonresponse rates. One general observation common to these papers is that for "core" data from the SIPP, the levels of item nonresponse are low.

In addition to the papers cited above, levels of item nonresponse can be found in the U.S. Bureau of the Census, Current Population Reports, Series P-70, Economic Characteristics of Households in the United States. Table 7 provides a summary of SIPP item nonresponse rates for each calendar quarter of 1984 compared to the March 1985 Current Population Survey.

Longitudinal Item Nonresponse

The concept of cross-sectional item nonresponse based on data obtained in one interview can be extended to a longitudinal concept that combines the nonresponse experience for successive interviews. This has been done for the first three observations in the 1984 SIPP; the results for a selected group of income types are shown in table 8. The rates in this table are based on the total number of persons reporting receipt of the specified income type at any time during the 12-month period. The first column shows the percent of all income recipients that reported amounts for all months during which the income source was received. The other columns indicate situations in which amounts were not reported in one or more, one or more but not all, and all months of recipiency. The right-most column showing the proportion of cases for which no income amount was reported indicates that only in a very small number of cases was no amount of information available. In other longitudinal nonresponse situations some reported values are present and thus available for use in filling in the missing responses when developing annual summaries from the monthly data.

Table 7. -- Item Nonresponse Rates for SIPP and March 1985 CPS, for Selected Income Types

Income Type	SIPP 1984 1st Quarter Monthly Average	SIPP 1984 2nd Quarter Monthly Average	SIPP 1984 3rd Quarter Monthly Average	SIPP 1984 4th Quarter Monthly Average	March 1985 CPS
Wage and Salary	7.2	7.5	7.5	7.6	18.9
Self-Employment Income	16.8	16.2	16.0	16.1	26.5
Federal Supplement Security Income	7.6	8.4	8.1	8.4	19.9
Social Security Income	10.8	11.6	11.7	12.3	21.9
Aid to Families with Dependent Children	6.1	6.9	6.5	5.5	16.0
Unemployment Compensation	10.1	13.6	10.4	12.7	21.8
Company or Union Pension	13.9	14.0	12.8	14.7	24.0
Food Stamp Allotment	5.2	6.3	6.7	6.6	13.7
Veterans' Compensation or Pension	11.3	11.2	11.9	13.5	18.3

1. U.S. Bureau of the Census, Current Population Reports, Series P-70, No. 3, Economic Characteristics of Households in the United States: First Quarter 1984, U.S. Government Printing Office, Washington, D.C. 1985

2. U.S. Bureau of the Census, Current Population Reports, Series P-70, Economic Characteristics of Households in the United States: Second Quarter 1984, U.S. Government Printing Office, Washington, D.C. 1985

3. U.S. Bureau of the Census, Current Population Reports, Series P-70, Economic Characteristics of Households in the United States: Third Quarter 1984, U.S. Government Printing Office, Washington, D.C. 1985

4. U.S. Bureau of the Census, Current Population Reports, Series P-70, Economic Characteristics of Households in the United States: Fourth Quarter 1984, U.S. Government Printing Office, Washington, D.C. 1986

TABLE 8. -- Longitudinal Item Nonresponse Rates for Amounts of Selected Income Types: 1984 SIPP Panel 12-Month Summary

(percent)

INCOME TYPE	ALL AMOUNTS REPORTED	ONE OR MORE AMOUNTS NOT REPORTED	ONE OR MORE BUT NOT ALL AMOUNTS NOT REPORTED	NO AMOUNTS REPORTED
Hourly Wage Rate . . .	83.0	17.0	9.0	8.0
Social Security. . . .	82.8	17.2	13.1	4.1
Private Pension. . . .	78.8	21.8	13.6	8.2
AFDC	91.0	9.0	5.6	3.4
Food Stamps.	91.9	8.1	6.2	1.9
Unemployment Compensation.	87.9	12.1	4.0	8.0
Federal SSI.	88.0	12.0	7.6	4.4

NOTE: These rates are based on the total number of persons with recipiency in one or more of the 12-months. Also these rates do not reflect imputations made to type Z person noninterviews.

The treatment of item nonresponse in a longitudinal manner has been a topic of substantial interest for the SIPP (Kalton and Lepkowski, 1983). Some work has occurred since the first year of SIPP which lays the foundation for future evaluations of SIPP longitudinal imputation systems.

An investigation of the feasibility of using model-based imputations has been conducted (Huggins and Weidman, 1986a, Huggins and Weidman, 1986b). Models which impute missing response patterns based on the frequency distribution of response patterns have also been investigated (Samuhel, Huggins, 1984; Huggins, Samuhel, and Weidman, 1985), and several imputation procedures for continuous data were compared in a small simulation study (Huggins, 1986).

Evaluation of Survey Reports of Income

Most evaluation of the quality of income and data derived from a household survey include comparisons of the survey estimates with comparable estimates derived from independent sources such as administrative records, tax data, etc. In this section we consider comparisons of income amounts. Comparisons of SIPP income estimates with independently derived estimates are shown in table 9 for a selected group of income types. These comparisons have been made based on the aggregate income received by the population. These comparisons indicate some variation in the ratios within a year between different income surveys. Typically the estimates of income from the survey fall short of those derived from independent sources that are, in general, considered benchmarks for measuring accuracy. The shortfall in the SIPP estimates for monthly figures are, in most cases, less than the CPS shortfalls for annual amounts. The SIPP estimate for wages and salary income, however, has a slightly larger shortfall than the CPS.

Table 9. -- Comparisons of Estimated Numbers of Income Recipients and Estimated Aggregate Income Amounts Received for Selected Income Types: SIPP vs Independently Derived Estimates vs the Current Population Survey

	SIPP As A Percent of The Independent Estimates of Monthly Average Recipients for Selected Income Types by Quarter	SIPP As A Percent of The Independent Estimates of Aggregate Income Amounts Received for Selected Income Types by Quarter	CPS (1983) as a Percent of the Independent Estimate Aggregate Income Amounts Received
Wage and Salary			
3rd Quarter 1983	(x)	95.0	99.0
4th Quarter 1983	---	94.3	
1st Quarter 1984	---	93.2	
2nd Quarter 1984	---	94.4	
3rd Quarter 1984	---	95.2	
4th Quarter 1984	---	94.5	
Federal Supplemental Security			
3rd Quarter 1983	92.0	89.8	84.9
4th Quarter 1983	91.3	93.5	
1st Quarter 1984	94.8	96.4	
2nd Quarter 1984	98.7	97.4	
3rd Quarter 1984	98.3	98.6	
4th Quarter 1984	98.1	99.2	
Social Security			
3rd Quarter 1983	99.2	99.6	91.7
4th Quarter 1983	96.3	100.6	
1st Quarter 1984	97.3	100.5	
2nd Quarter 1984	97.7	101.1	
3rd Quarter 1984	97.5	101.3	
4th Quarter 1984	97.5	101.6	
Aid to Families With Dependent Children 1/			
3rd Quarter 1983	78.5	76.2	76.0
4th Quarter 1983	79.2	78.5	
1st Quarter 1984	84.6	85.3	
2nd Quarter 1984	86.0	86.0	
3rd Quarter 1984	82.0	80.2	
4th Quarter 1984	80.7	78.8	
Unemployment Compensation			
3rd Quarter 1983	102.6	100.9	75.5
4th Quarter 1983	103.4	106.8	
1st Quarter 1984	82.6	85.2	
2nd Quarter 1984	82.5	83.1	
3rd Quarter 1984	78.5	80.3	
4th Quarter 1984	95.1	100.9	
Food Stamps			
3rd Quarter 1983	89.5	90.1	71.2
4th Quarter 1983	91.1	83.1	
1st Quarter 1984	90.8	85.2	
2nd Quarter 1984	90.5	86.2	
3rd Quarter 1984	90.3	84.6	
4th Quarter 1984	91.7	83.6	
Veterans' Compensation or Pension			
3rd Quarter 1983	89.2	78.9	63.3
4th Quarter 1983	89.7	79.9	
1st Quarter 1984	90.6	78.0	
2nd Quarter 1984	90.8	74.5	
3rd Quarter 1984	89.8	76.3	
4th Quarter 1984	93.3	79.7	

1/ On Table D-2, the amount excludes dependent covered by payments.

(x) Not applicable.

SOURCE: U.S. Bureau of the Census, Current Population Reports, Series P-70, Nos. 1, 2, 3, 4, and 5 Economic Characteristics of Households in the United States

Evaluation of Labor Force Estimates

One way to evaluate labor force estimates in SIPP is to compare them to those from the CPS. Even though such comparisons are not definitive, some preliminary judgements can be made as to how well SIPP is measuring labor force activity over time, at least in the aggregate.

In the following, trends in SIPP labor force estimates between the third quarter of 1983 and the first quarter of 1985 have been compared to those from the CPS. The SIPP and CPS labor force data used in these comparisons have not been seasonally adjusted. More importantly, the data have not been adjusted for conceptual and measurement differences between the two surveys--and differences do exist (some of which are discussed below). Perhaps the two most significant ones involve coverage and the reference period. In SIPP, persons from farm households are excluded in the data and the reference period is a full month; in the CPS, the data relate to all households and the reference period is the week containing the 12th of the month.

The first comparison in figure 3 concerns employment. The SIPP data represent persons with jobs during a month; the CPS data reflect persons who worked during the reference week (or had a job from which they were absent) of a month. The trends in each series sketch out similar patterns, although there are a couple of divergences. Between the third quarter of 1983 and first quarter of 1984, employment as measured by the CPS declined while the SIPP estimate of persons with jobs rose. In addition, the employment drop registered in the CPS between the fourth quarter of 1984 and first quarter of 1985 was greater than that recorded in SIPP.

Figure 4 contains a comparison of persons looking for work or on layoff according to SIPP and persons who were unemployed according to the CPS. The SIPP estimates are much higher than the CPS estimates, but the overall trends in both series tend to parallel one another. The most significant trend differences, however, occurred between the fourth quarter of 1983 and first quarter of 1984 when CPS unemployment increased while the comparable SIPP estimate declined.

A last comparison displayed in figure 5 shows persons with no labor force activity from SIPP and persons classified as not in the labor force from the CPS. Again, level differences are apparent, but the trends are very much alike. Two exceptions that should be noted occurred during the second and third quarters of 1984 and between the last quarter in 1984 and first quarter in 1985.

An investigation into the "level" differences in the labor force data from both surveys was conducted jointly by the Census Bureau and Bureau of Labor Statistics (Ryscavage and Bregger, 1985). Labor force estimates from these surveys were compared <u>after</u> two adjustments were made to the SIPP data: persons living in farm households were included in the SIPP estimates and the CPS reference period was replicated in the SIPP. It was found that the SIPP estimate of persons with jobs was slightly lower than

the CPS estimate of employment, while the unemployment estimate in SIPP continued to be about 16 percent higher (before the adjustment, the difference was 28 percent). The discrepancy in the "not in the labor force" estimates also narrowed after the adjustment. Recall problems, questionnaire differences, and other unique features of each survey were suggested as possible causes of the remaining differences in the labor force estimates.

SIPP Gross Flow Data

Analysis of program data on a month-to-month basis in SIPP's predecessor program, the Income Survey Development Program, revealed a tendency for reported program turnover to occur between waves of interviewing more often than within the wave (Moore and Kasprzyk, 1984). Further analysis on this topic was deferred until the availability of three observations from the SIPP last summer. The analysis (Burkhead and Coder, 1985) which covered month-to-month changes in recipiency of income benefit amounts for a 12-month period focussed on changes occurring between the last month of one reference period and the first months of the succeeding reference period (between months 4 and 5 and between months 8 and 9 in sample.) Not surprising, the results using SIPP data are similiar to the ISDP results, where an uneven pattern of change is observed and that this pattern is clearly associated with the interviewing scheme. Gross changes are significantly higher between the last month of one reference period and the first month of the next; further work is necessary, however. The main causes for the problem seem to be questionnaire wording/design, respondent recall error, and the interaction between these two factors.

The concern about this topic has increased since the first data were collected in the SIPP, and several research projects have begun to study the potential seriousness of biases in gross flow estimates due to response error. Staff in the Statistical Methods Division have developed a model to estimate biases in gross flow estimates that result from response error. They have developed a methodology for estimating the parameters of the model using SIPP response error rates and the ratios of within-wave and between-wave gross flow estimates. SIPP reinterview program data are being evaluated for the possible use in estimating response error rates. Using this model and methodology the bias for many hypothetical gross flow estimates can be computed. Preliminary results will be presented at the 1986 meeting of the American Statistical Association (Hubble and Judkins, 1986). Alternative models will also be explored when SIPP longitudinal data file is available.

Staff in the Statistical Research Division are also involved in an effort to try to determine the causes for bias in SIPP gross flow estimates. They intend to do exploratory analysis of SIPP data to determine what variables are related to over/under reporting and if imputed responses are responsible for a disproportionate number of between wave transitions. They also will estimate models for the effects of error sources on the number of transitions and the probability of transitions. Preliminary results of this work will be presented at the 1986 meeting of the American Statistical Association (Weidman, 1986).

Figure 3 SIPP-CPS Labor Force Comparisons, I

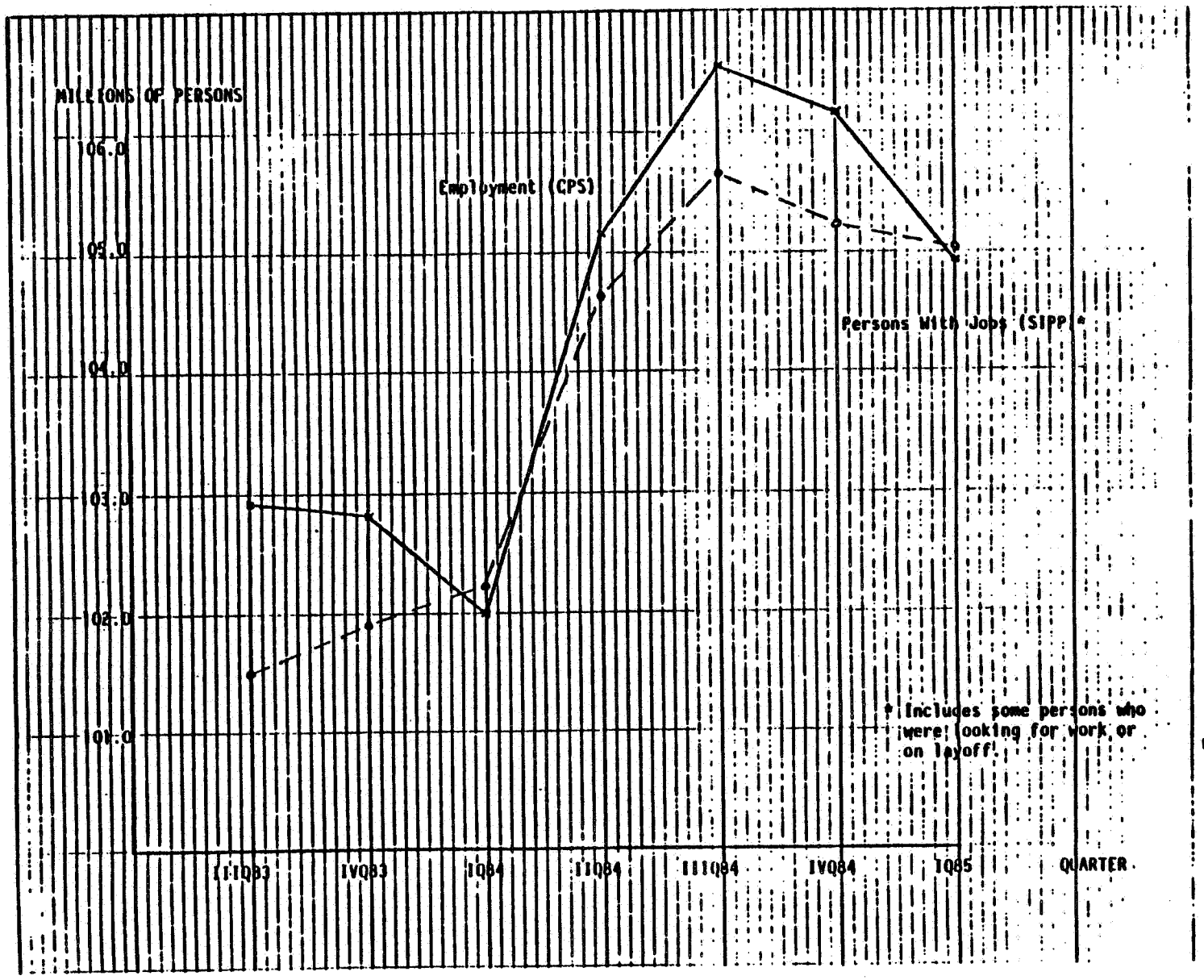


Figure 4 SIPP-CPS Labor Force Comparisons, II

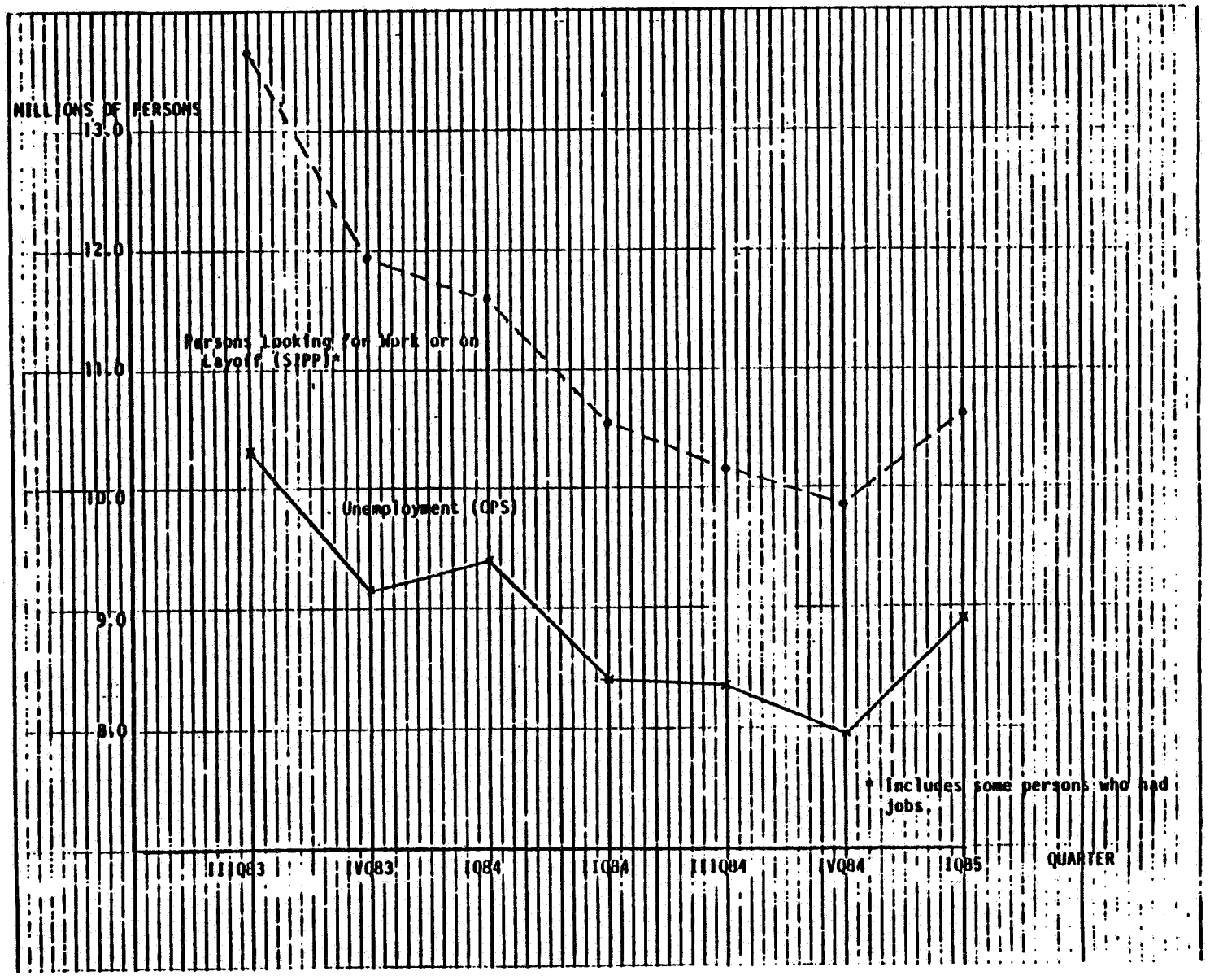
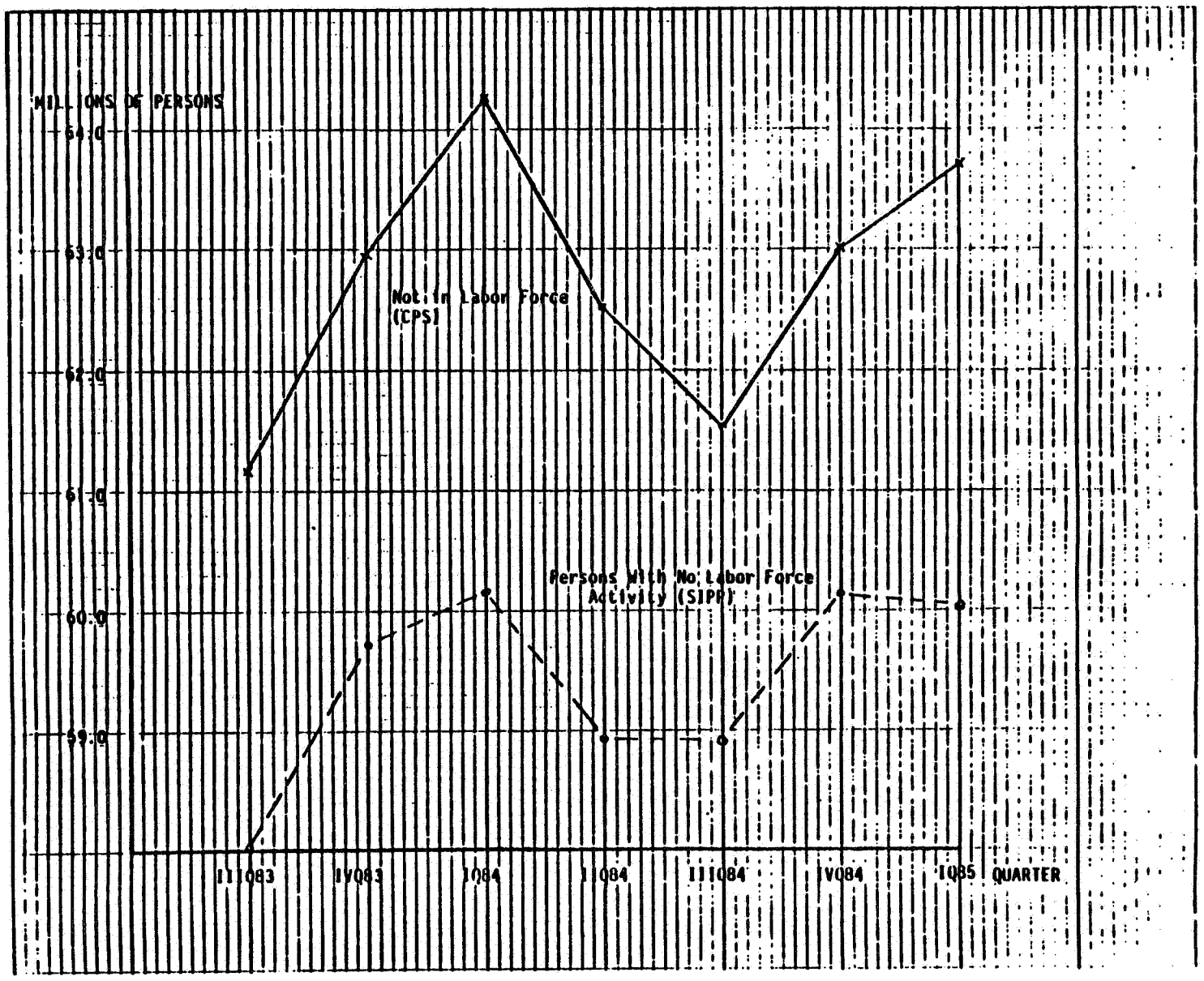


Figure 5 SIPP-CPS Labor Force Comparisons, III



Reinterview Program

Another aspect of evaluation and one important for maintaining control over the data collection effort is the reinterview program. The reinterview program for SIPP was designed specifically as a quality control of the interviewer's work. Each month one-sixth of the interviewers are selected for reinterview. One-third, but no more than eight units, are chosen from each selected interviewer's assignment. During the first wave of the 1984 and 1985 panels all new SIPP interviewers are selected once for reinterview, but only three sample housing units are chosen from their assignments. This is done in order to obtain an early reading on the performance of the SIPP interviewer staff.

Results are available from the 1984 Panel Reinterview Program for Waves 2 through 4 (St. Clair, 1985) and can be summarized as follows:

1. The gross and net error rates in household composition (.19 percent and -.08 percent, respectively) and the noninterview misclassification rate (1.8 percent) were below the median rates in the Consumer Expenditure, National Crime and Health Interview Surveys.

2. The percent of interviewers failing the sample unit, household composition, and content checks were all below 2 percent.

3. The only areas of concern were two content item groups: health insurance, and assets. Interviewers accounted for almost one-half of the differences detected in these item groups.

For the time period of December 1983 through August 1985, five SIPP interviewers were caught fabricating interviews.

V. PLANNED EVALUATION OF SIPP DATA

Census Bureau staff will continue evaluations of data quality the same ways indicated above. Often the evaluation will be written in the context of concern about a substantive issue. For example, a recent paper on SIPP labor force transitions (Ryscavage and Short, 1985), prepared for presentation at the annual meeting of the American Economic Association, has done much to better the understanding of the use of the "employment status recode" in SIPP and its relationship to the CPS. In the future, however, an Evaluation Steering Committee will be formed to identify priorities in the evaluation of SIPP data and ensure that the projects identified are completed. Thus, we expect a continuing staff involvement and committment to evaluation.

Three other groups of people will also contribute to SIPP data evaluation: 1) the Center for Survey Methods Research; 2) participants in the American Statistical Association/Census Bureau/National Science Foundation Fellowship program; and 3) other Federal users of SIPP data. A discussion of their projects follows:

Center for Survey Research Methods Research: Multi-State Multi-Program Record Check Study

A record check project was initiated to investigate response quality issues in SIPP through a case-by-case comparison of SIPP data with administrative record information. The principal investigator for the project is on the staff of the Center for Survey Methods Research under the Associate Director for Statistical Standards and Methodology. The research questions to be addressed in the project include: 1) the quality of recipiency status and benefit amount reporting for a variety of state and Federally administered transfer programs; 2) the effects of recall period length on report quality; 3) the nonexperimental effect of self/proxy respondent status on report quality; 4) the extent of misclassification errors; 5) between wave turnover effects; and 6) demographic correlates of report quality.

Four State administered programs (Aid to Families with Dependent Children, food stamps, unemployment insurance, and worker's compensation) and six Federally administered programs (Civil Service Retirement, Old Age Surviors and Disability Insurance, Pell Grants, Supplemental Security Income, and Veterans Pension and Compensation) have been selected for inclusion in the study. Four States were selected as study sites, based on the following criteria: 1) a high-quality computerized record system for all programs of interest; 2) no confidentiality restriction on sharing individually identifiable data with the Census Bureau for research purposes; and 3) a substantial SIPP sample. The States selected are Florida, New York, Pennsylvania, and Wisconsin. Work on this project has proceeded slowly largely because acquiring the administrative and legal arrangements to use administrative data from States and Federal agencies is very time consuming. Approximately 80 percent of the data files have now been received and our analysis plan is now available (Moore, 1986).

ASA/NSF/Census Bureau Fellowship Program

Analysts who were awarded Fellowships in 1986 as part of the American Statistical Association/National Science Foundation/Census Bureau Fellowship Program will also participate in the evaluation of SIPP data. Fellowship awards for SIPP research were made to:

1. Pat Doyle (Mathematica Policy Research) to study the potential for SIPP to analyze serial program participation over a 12-month period concentrating on transitions among multiple benefit categories, the sequence of these transitions, and the length of time between transitions. Ms. Doyle also plans to examine the impact of longitudinal unit definitions, and longitudinal edit and imputation procedures on measures of serial multiple program participation, thereby providing the Census Bureau an evaluation of the preliminary longitudinal data product.

2. Martha Hill (University of Michigan) to study the sensitivity of annual measures of economic status to the treatment of household composition change over the year. Dr. Hill's analysis intends

VI. ANALYSIS AND PRODUCTS AT THE CENSUS BUREAU

Reports and Files

Analysis of SIPP data proceeds along several dimensions at the Census Bureau. The Current Population Reports, Economic Characteristics of Households in the United States, Series P-70, is the principal vehicle for the release of SIPP data in printed form; up to this time these reports have presented average monthly labor force activity, income, and program participation statistics (U.S. Bureau of the Census, 1984, 1985a, 1985b, 1985c, 1985d, 1986). In the near future, the reports will emphasize analyses using data collected in the topical modules.

Public-use microdata files, consisting of unaggregated records for individual survey respondents, are released to the public. Several files containing core data collected in a wave are now available. The wave files are available for purchase by the public in both a relational and rectangular format. The relational structure is fairly complex and difficult for nonprogrammers to use. There are eight types of records: sample unit, household, family, person, wage and salary, self-employment, asset income, and other income sources. The relationships between the records are expressed by pointers on each record.

Some users, particularly those who want to use statistical software packages, prefer the rectangular structure. A rectangular file has one type of record of consistent length and fixed format throughout. Each logical record for a sample person includes information on the household and family of the person during each month of the reference period, as well as characteristics of the person and each source of income received. The relational and rectangular wave files are fully edited, imputed, and weighted for use in cross-sectional analyses.

Despite a set of public-use products, it seems obvious that the market for SIPP products is still evolving. The kinds of reports and files being produced today are not necessarily the products of the future; these products will be determined by the user community response to the current products. A discussion of a typology of reports and analyses possible using SIPP data as well as some ideas about the creation of microdata files from SIPP is available in attachment 5. This attachment provides a different way of thinking about SIPP products; as we plan for future products, it will need to be discussed within the SIPP user community.

Attachment 6 provides a listing of the written products which are now available. It includes not only official reports (Series P-70) but also compilations of papers presented at conferences and SIPP Working Papers.

Finally, census staff have been active in presenting substantive and methodological findings at professional meetings; in particular, papers have been or will be presented on SIPP at the following professional meetings: 1) American Statistical Association (1983, 1984, 1985, and 1986); 2) Population Association of America (1986); 3) American

to focus on comparisons of SIPP, CPS, and Panel Study of Income Dynamics (PSID) estimates of household income for 1984. The comparison will attempt to distinguish between the influence of the method of treating household composition change and the influence of other differences in design. As a result of this work, census staff should have a better understanding of SIPP income data and its relationship to other income surveys.

3. Daniel Hill (University of Michigan) to use SIPP data to investigate the effects of length of recall, and to a lesser extent conditioning, on the quality of such measures as monthly earnings, transfer program participation and income, labor supply, and unemployment. This research will further our understanding of the reporting process and help in understanding and reducing the differences in the SIPP data between month-to-month changes measured within and between interviewing waves.

4. Heidi Hartmann (National Academy of Sciences) to use SIPP data to 1) identify the economic correlates of increased nonmarriage, particularly the increased number of female headed households with children, and to explain the differing rates between Black and White women, and 2) to examine the correlates of success of female-headed households where success is defined as the ability to support oneself and children above the poverty level, and, in particular to investigate whether the poverty of female-headed households is associated with nonmarriage.

In each case, the Fellow's research will improve the understanding of SIPP data through a systematic evaluation of the quality of the SIPP data.

Federal Users of SIPP Data

The principal users of SIPP data will be analysts at Federal agencies who must respond to agency requests for information on specific policy issues. Census Bureau staff has maintained contact with a number of analysts and has encouraged their use and reporting of SIPP data at professional forums. Thus, staff at the Food and Nutrition Service (Department of Agriculture) were asked to made a substantive contribution to the Census Bureau's Second Annual Research Conference. FNS responded by developing a paper which makes preliminary comparisons between the Agency's administrative records for food stamps and the SIPP (Dalrymple and Carlson, 1986). Three other Federal agencies are in the process of developing research papers for presentation at an invited paper session at the 1986 meeting of the American Statistical Association. An analyst from the Congressional Budget Office will discuss poverty rates and program participation in the SIPP; an economist from the Department of Health and Human Services will discuss multiple transfer program participation and its effect on the poverty rate; and an analyst at the Congressional Research Service will examine the labor force participation of women.

While these evaluation activities are outside of the work scope of Census Bureau staff, they are included here because the studies were suggested by census staff with the expectation that a better understanding of the quality of the SIPP data will result.

Association of Public Opinion Research (1986); 4) American Economic Association (1985).

While much activity exists in report and research paper preparation, it is quite clear that many audiences, including Congress, heads of other government agencies and their immediate staff or nongovernmental agencies, are currently inadequately served by Census Bureau publications (Vaiana, 1985). Thus, census staff is now exploring the possibility of developing publications in a format different from our current reports. The type of report being considered would be short, timely, more issue oriented, and nontechnical. At this time a draft of a report is circulating for comment, and the approach offered by this product is under discussion.

Longitudinal Research File

A research file of SIPP observations at three points in time is being created for staff use in order to better understand the multiwave data collected in SIPP. The variables, file structure, and the edit and imputation specifications for income variables are now being defined.

Work on the longitudinal edit and imputation system for demographic data has also begun. The general plan for the longitudinal demographic edit involves six steps. First, cross-sectionally edited data are reviewed to identify longitudinal inconsistencies in the following variables: age, sex, race, marital status, relationship to reference person, person number of parent, person number of spouse, person number of guardian, reason for entering/leaving, and type of living quarters. Second, preliminary edit/imputation specifications are developed for each variable to resolve inconsistencies and other problems with cross-sectional definitions when viewed longitudinally. Third, results of preliminary specifications are tested. Fourth, preliminary specifications are revised based on assessment of empirical results derived using preliminary specifications. Fifth, results from modified specifications are reviewed to ascertain whether cross-sectional consistency was maintained through longitudinal edit, and longitudinal edit is revised as necessary. Sixth, specifications are implemented in computer software.

Certain demographic variables collected in the SIPP have been carefully controlled since the beginning of the survey. As a result only a very minor number of inconsistencies from wave-to-wave exist in the demographic data. Table 10A, B, and C illustrate how small this problem is. Development and implementation of income and demographic longitudinal edit and imputation specifications, and weighting specifications will continue through the summer.

TABLE 10A, B, and C.

Race Change in the 1984 SIPP Panel

Table A -- Inconsistencies in 1984 Panel Demographic Data: Wave-to-Wave Race Change

From Wave	To Wave	White to Black	White to American Indian, Eskimo, Aleut	White to Asian or Pacific Islander	Black to White	Black to Asian or Pacific Islander	American Indian, Eskimo, Aleut to White	Asian or Pacific Islanders to White	Refused to White	TOTAL
1	2	31	0	1	2	0	10	2	0	46
2	3	16	6	1	2	0	1	0	1	27
3	4	31	7	0	3	1	0	2	0	44
4	5	13	10	5	4	0	2	5	0	39
TOTAL:		91	23	7	11	1	13	9	1	156

Sex Change in the 1984 SIPP Panel

Table B -- Inconsistencies in 1984 Panel Demographic Data: Wave-to-Wave Sex Change

From Wave	To Wave	Male to Female	Female to Male	Total
1	2	22	12	34
2	3	16	18	34
3	4	6	16	22
4	5	20	13	33
TOTAL		64	59	123

Age Change in the 1984 SIPP Panel

Table C -- Inconsistencies in 1984 Panel Demographic Data: Wave-to-Wave Age Change

		AGE INCREASED			AGE DECREASED			
From Wave	To Wave	2 - 5 years	6 -10 years	11 or more years	2 - 5 years	6 -10 years	11 or more years	Total
1	2	33	8	3	24	5	9	82
2	3	50	13	19	88	10	9	189
3	4	51	8	3	65	5	1	133
4	5	14	8	2	51	3	1	79
TOTAL		148	37	27	228	23	20	483

These tables were developed from administrative records. The uncertain data quality of those administrative records may result in these numbers differing from those developed from public-use data files. Adult sample size (15+ years old) in Wave 1 was 43,664. Children sample size (less than 15 years old) in Wave 1 was 10,071.

Longitudinal Concepts

Household and family level analysis in a longitudinal survey is complicated by the fact that the composition of households and families can change over time since original sample persons leave to join other households or families, or to set up new ones. The principal issue is the development of definitions of households and families which account for survey measurements at two or more points in time and which do not create serious conflicts with the traditonal cross-sectional household and family constructs. The definition has been refined through numerous internal discussions (McMillen and Herriot, 1985; Citro, 1985), and as a result the Population Division, in consultation with other divisions, has developed and provisionally adopted for research use with SIPP a longitudinal household concept. Software to implement this concept has been written, providing the capability for creating SIPP microdata files that can be used to study the formation, continuation, and dissolution of households through time.

The major components of the Bureau's longitudinal household concept can be summarized as rules which are applied to households 2 consecutive months at a time. Rule 1 is that a family household maintained by a married couple continues as long as the couple maintains a household. Rule 2 is that a family household not maintained by a married couple continues as long as at least one family member continues to live with the householder. Rule 3 pertains to cases in which a married-couple household is either preceded or followed by a situation in which the husband or wife in the married-couple household was maintaining a family household separate from the spouse. Under rule 3, the married-couple household is continuous with one of the other family households, if a majority of family members in the married-couple household are also present as a majority in the other household. Rule 4 is that a oneperson nonfamily household continues as long as the householder maintains such a household. Rule 5 is that a multiperson nonfamily household continues as long as the householder maintains such a household. Rule 6 is that a household continues, if it changes type from nonfamily to family because two unmarried persons living together have become married to each other. In cases of conflict between rules 5 and 6, rule 6 takes precedence.

The Bureau's longitudinal household concept can be seen as a hybrid of several other possible concepts. Three of the most reasonable alternatives would define a household as continuing under the following conditions. First, the reference person definition would view a household as continuing over time if it has the same reference person or householder. Second, the principal person definition would view a household as continuing over time if it has the same principal person, where the principal person is the householder except in married-couple households where the principal person is the wife. Third, the family type definition would view a household as continuing over time if it has the same reference person and if it is the same family type, where family types are married-couple household; other family household, male householder; other family household, female householder; nonfamily household, male householder; nonfamily household, female householder.

In order to provide evidence regarding the consequences for basic analyses of differences between the Bureau's concept and major alternative concepts, research is currently underway. This research will produce results pertaining to the existence and magnitude of differences across five different longitudinal household concepts with regard to: (a) household type, (b) household duration, (c) household composition, and (d) low-income status by household type and duration (Citro, Hernandez, and Herriot, 1986).

Longitudinal Processing

In order to facilitate the processing of SIPP data in general and longitudinal processing in particular, the Census Bureau is proceeding along two paths: 1) by developing software for creating, extracting, and analyzing longitudinal data; and 2) by researching alternative methods of accessing large complex data sets.

One of the analysis problems presented by SIPP is that many of the statistics that analysts want to calculate from SIPP data are not options in any tabulation or statistical software. Within the Census Bureau an effort has begun to develop software to calculate spell characteristics such as duration, type of censoring, and beginning and ending dates of spells. In addition, this software is designed to retrieve characteristics of individuals during spells in a state such as modal labor force status or mean income during receipt of food stamps. Outside the Census Bureau, we have held several discussions with the OSIRIS Development Group at the University of Michigan as they have developed similar capabilities within their statistical software package.

Census Bureau staff has also continued research into alternative methods of accessing large complex data sets. Following on internal experimentation with SIPP and the Scientific Information Retrieval (SIR) data base management system, we are evaluating the feasibility of using RAPID to access SIPP data. RAPID is a data base management system developed by Statistics Canada for processing Canadian Census and survey data. Two features of RAPID make it attractive for SIPP processing. First, RAPID works with a transposed data structure which improves efficiency. Second, RAPID has interfaces with SPSS, SAS, and TPL. These user friendly interfaces are central focus of our investigation. We are focusing on the ease and efficiency with which analysts can access a large complex data base.

VII. ACTIVITIES RELATED TO USER NEEDS

Census staff has been active at a number of forums where issues in the use of SIPP data were actively discussed. The Census Bureau financially supported and participated in a conference cosponsored by the Social Science Research Council and the National Science Foundation. This conference provided the research community an opportunity to provide constructive criticism of the SIPP design and questionnaire to the Census Bureau. The proceedings of this conference, including recommendations, are available in a special issue of the Journal of Economic and Social Measurement (David, 1985).

Census staff has also participated in a conference held by the Social Science Research Council under a grant from the Ford Foundation. The purpose of the conference was to draw on participants' experiences with many different data sets and their substantive and methodological knowledge in order to propose a feasible research agenda for the SIPP to the research community. In this way the conference would contribute to the Ford Foundation's policy-focussed "Project on Social Welfare and the American Future" by demonstrating the potential of SIPP as an instrument for social scientists whose research questions are fundamental to policy analysis. A report of this conference is in preparation by Social Science Research Council Staff.

Census staff has also been active in conducting workshops for users of SIPP data at professional association meetings such as the meetings of the American Sociological Association (1984, 1985, 1986), the Population Association of America (1985), Southern Regional Demographic Group (1984), and the American Economic Association (1985), as well as to interested analysts in the Washington D.C. area (1984). A more extensive one-week workshop was held as part of the Inter-University Consortium for Political and Social Research (ICPSR) Summer Training Program in Quantitative Methods of Social Research in July 1985. Census staff will participate in another SIPP Workshop sponsored by the ICPSR during the summer of 1986.

As more analysts have begun to use SIPP data, the need for a vehicle by which to exchange problems and solutions in using SIPP data has become apparent. In response to the need, census staff have organized and are chairing a Washington, D.C. SIPP Users' Group. This group intends to meet approximately every three weeks. Census Bureau presentations on the public-use file structure, variables, weights, and matching considerations are planned; analysts at other agencies will also speak concerning the problems observed in using SIPP data for their projects.

The issue of access to SIPP data is obviously critical to all potential users of SIPP data. Workshops and Users' Group chaired by Census Bureau staff help to allay doubts and fears about using the new data base. Census staff are also proceeding along three additional paths designed to help the user community.

First, another part of the examination of alternative methods of accessing SIPP data is our cooperation with the University of Wisconsin's NSF-funded project to develop a SIPP data center. Features of this project which are of particular interest include: 1) the Census Bureau provides the project with a complete set of internal memoranda, periodically updated, as part of the documentation process. The title, subject, and authors of those memoranda are entered into a data base management system (INGRES) for easy retrieval. That same data base management system is used to provide easy access to the complex structure of SIPP data, 2) Census Bureau personnel participate in the computer conference to discuss SIPP data processing, accessing data through INGRES from Suitland, and assist in the training of those wishing to use SIPP through the University of Wisconsin data center.

Second, the Census Bureau is currently conducting a thorough review of the technical documentation for SIPP. The goal of this review is to produce a single record layout and file description that can be used for Waves 2-9 of the 1984 Panel and Waves 2-8 of the 1985 Panel.

Third, a draft version of six chapters of a SIPP User's Guide has been written and reviewed. Three other chapters are being drafted. Those nine chapters will be edited and subject to a final review before publication. The primary complication in this process is the lack of staff time to incorporate suggested clarifications and improvements.

VIII. OTHER TOPICS

Statistical Estimation for Longitudinal Concepts

Weighting the longitudinal sample, especially for analytic units other than the individual, is a difficult area requiring a continuing statistician-analyst dialogue. Research on this topic has proceeded along two dimensions--longitudinal person estimation and longitudinal household (family or recipient unit) estimation. Since, at a minimum, the first SIPP longitudinal microdata products will be files of person-based information with household (family or recipient unit) as an attribute of the person, the early emphasis of the work is on the longitudinal person estimation. The work includes the calculation of selection probabilities to yield unbiased longitudinal estimates of individual characteristics and the use of controls in additional stages of estimation (Judkins, Hubble, Singh, McMillen, and Ernst, 1984). A refinement of this work and a description of the method proposed to produce longitudinal weights for person analysis covering the first three SIPP interviews will be available this summer (Kobilarcik and Singh, 1986).

The topic of longitudinal household (family or recipient unit) estimation is also under study. Some early approaches to this issue were presented at the 1984 meeting of the American Statistical Association (Ernst, Hubble, and Judkins, 1984). After additional study and tentative adoption of a longitudinal household definition, methodology was developed to obtain unbiased weights and an approach to weighting adjustments was outlined (Ernst, 1985; Ernst, 1986).

Composite Estimation

Composite estimation is a technique that combines estimates from the current and previous time periods with the goal of improving the precision of survey estimates by taking advantage of the correlations between response for the same analytic units at different time periods. Composite estimation is particularly effective when the correlations are high, which is likely to be the case for many important data items in SIPP. A preliminary review of types of composite estimators appropriate for the SIPP data structure is available (Chakrabarty, 1986), and research on this topic will continue this year.

Variance Estimation

Census staff has been concerned about the ability of users of SIPP data to develop variance estimates from half-sample codes available on the SIPP public use files. As a result, a review of generalized variance estimation software was conducted (Smith, 1984), and staff from the Statistical Research Division have been asked to develop a chapter in the SIPP User's Guide which would explain in detail how analysts can use the software packages and the data on the file to produce their own variance estimates.

Cross-sectional Estimation

This research has just begun and will focus on the impact of the current second-stage weighting procedure on the SIPP cross-sectional estimates. The procedure is very complex, and its effects have not been fully explored. The complexity primarily results from the requirement that the husband and wife in each married-couple family have the same weight. The effects of this requirement will be examined and the current method of meeting it will be compared with some alternative approaches. The work will include an analysis of the stability over time of the estimated number of individuals in specific demographic subgroups with respect to each weighting procedure under consideration.

SIPP Wealth Data

Considerable interest has arisen in the research community concerning the collection of wealth data in SIPP. Recognizing the household survey's limitations in the collection of such data, census staff has asked Dr. Courtenay Slater to evaluate the need for data on household asset holdings with emphasis on Federal government uses of the data. Dr. Slater will describe present sources other than SIPP for information on household assts and analyze the advantages for obtaining asset data through the SIPP, such as the opportunity to relate asset holdings to income and other household characteristics. She will also discuss the limitations of the SIPP and provide recommendations for possible design and content alterations for the improved collection of wealth data. As part of her work, Dr. Slater will also explore non-Federal uses of asset and income data from the SIPP, including their potential use by the private business sector for economic and market analysis. She will also recommend content changes to SIPP data products which would increase their usefulness to non-Federal users. Her work in these areas will be completed this summer.

Matching Activities

The SIPP data system has always been thought of as a combination of data from administrative records and household surveys. This reduces respondent burden by using other data sources for difficult-to-obtain information. Interview responses can be supplemented by information from program files such as the earnings and benefit records of the Social Security Administration. To make these linkages accurate, social security numbers (SSN)

are required for sample individuals. The SSN is obtained for each household member in SIPP and recorded on the control card. It is identified as a critical survey data item requiring completion to make the interviewers aware of its importance.

To facilitate the possibility of matches to other record systems, a SSN Validation Project has been designed which verifies social security numbers collected from survey respondents, and obtains SSN's for respondents who could not provide one (excluding refusals).

Two files for original (100 level) sample persons in the 1984 Panel were created for use in matching SIPP data to administrative record systems:

1. The Validated SSN Link File, which contains the validated SSN's and corresponding demographic characteristics for 44,172 Wave 1 1984 Panel SIPP respondents; and

2. The Unvalidated SSN File, which contains 7,742 unvalidated cases (most of which are children) along with the corresponding demographic characteristics and an explanation of why each case was invalid.

200 and above level persons in the 1984 Panel will have their SSN's validated with 100-level persons in later SIPP Panels on a flow basis.

At the present time, the only match to administrative records which has taken place is a match to the extract of tax returns which the Bureau has in-house. The Social Security Administration has expressed interest in matching SIPP data with the Master Beneficiary Record and Supplemental Security Record. Census staff have been working with them to resolve technical and legal issues.

Finally, a working group has developed a research plan and methodology proposal in which SIPP demographic data would be merged with economic data from census files to form a microdata base for both individuals and the firms in which they work. This project includes: a) identification of the study areas and issues to which the data base would be applied; b) determining the availability, coverage, and quality of economic data in various census files; c) specifying the demographic and economic data contained in the new data base; and d) resolving methodological problems encountered in developing a microfile containing worker and firm data. (Haber, Ryscavage, Sater, and Valdisera, 1984; Haber, 1985; Sater, 1985.)

Other Analysis (Not covered elsewhere)

Two other topics, not discussed elsewhere, should be mentioned. First, SIPP data can be used to address a wide variety of migration topics by traditional cross-sectional analysis as well as the survey's longitudinal design which provides a natural source of geographical mobility data because individuals are followed when they move to a new residence. A review of analyses of migration from previous panel surveys and an assessment of how SIPP can further our understanding of geographic mobility processes has been completed (Dahmann, 1984; Dahmann, 1986).

Second, a demonstration study exploring the adequacy of the SIPP in using hazards models in analyzing SIPP labor force transitions has been completed (Short and Woodrow, 1985) and further work is planned. This year staff from the Statistical Research Division will consult with Population Division staff to develop statistically valid methods of estimation of model parameters for data from the complex SIPP design.

Variance Reduction through Post-stratification

Recent new statistics on the effect of sample reductions on the variance and our ability to measure changes in differences in the number of statistics have created serious concerns. These concerns have caused us to increase our exploration of ways to reduce the variance. One approach is through the use of administrative records for post-stratification. Currently, cross-section estimation procedures for SIPP make use of a second-stage adjustment to increase the precision of estimates by ratio adjusting collection month and reference month estimates to CPS March type population estimates. However, the Census Bureau has access to some Internal Revenue Service and Social Security Administration files which can be used to produce detailed age, race, and sex distribution of adjusted gross income. The issue, which we have just begun to explore, is how these administrative data can be used for post-stratification to improve estimates of mean and median personal and household income as well as the estimates of the deciles of the personal and household income distribution. Furthermore, a basic question which will be considered is how much reduction in variances of these estimates can be achieved through such a procedure. These issues will be researched during the next 6 months. Further information concerning this topic will be provided at the meeting.

REFERENCES

BAILEY, L., D. CHAPMAN, and D. KASPRZYK (1985).
Nonresponse Adjustment Procedures at the Census Bureau: A Review. Proceedings of the U.S. Bureau of the Census First Annual Research Conference.

BURKHEAD, D. and J. CODER (1985).
Gross Changes in Income Recipiency from the Survey of Income and Program Participation. Proceedings of the Survey of Research Methods Section, American Statistical Association, Washington, D.C.

CHAKRABARTY, R. (1986).
Composite Estimation for SIPP: A Preliminary Report. Forthcoming SIPP Working Paper.

CITRO, C. (1985).
Alternative Definitions of Longitudinal Households in the Income Survey Development Program: Implications for Annual Statistics. Proceedings of the Survey of Research Methods Section, American Statistical Association, Washington, D.C.

CITRO, C., D. HERNANDEZ, and R. HERRIOT (1986).
Longitudinal Household Concepts in SIPP: Preliminary Results. Proceedings of the U.S. Bureau of the Census Second Annual Research Conference.

CODER, J. and A. FELDMAN (1985).
Early Indications of Item Nonresponse in SIPP. Proceedings of the Survey of Research Methods Section, American Statistical Association, Washington, D.C.

DAHMANN, D. (1984).
Panel Surveys as a Source of Migration Data. Proceedings of the Social Statistics Section, American Statistical Association, Washington, D.C.

DAHMANN, D. (1986).
Geographical Mobility Research with Panel Data. The Professional Geographer, forthcoming.

DALRYMPLE, R. and S. CARLSON (1986).
Food Stamp Participation: A Comparison of SIPP with Administrative Records. Proceedings of the U.S. Bureau of the Census Second Annual Research Conference.

DAVID, M., editor (1985).
Special Issue: Survey of Income and Program Participation. Journal of Economic and Social Measurement, Vol. 13, No. 3-4.

ERNST, L. (1985).
The Impact of Conditions in a Longitudinal Household Definition on the Ability to Produce Unbiased Longitudinal Estimates in SIPP. U.S. Bureau of the Census report.

ERNST, L. (1986).
SIPP Longitudinal Household Estimation for the Proposed Longitudinal Definition. Forthcoming SIPP Working Paper.

ERNST, L., D. HUBBLE, and D. JUDKINS (1984).
Longitudinal Family and Household Estimation in SIPP. Proceedings of the Survey of Research Methods Section, American Statistical Association, Washington, D.C.

HABER, S. (1985).
Applications of a Matched File Linking the Bureau of the Census Survey of Income and Program Participation and Economic Data. SIPP Working Paper Series No. 8502, U.S. Bureau of the Census.

HABER, S., P. RYSCAVAGE, D. SATER and V. VALDISERA (1984).
Matching Economic Data to the Survey of Income and Program Participation: A Pilot Study. Proceedings of the Social Statistics Section, American Statistical Association, Washington, D.C.

HOLT, M. (1985).
Preliminary Report on Training Materials for the Survey of Income and Program Participation. Paper prepared for Population Division, U.S. Bureau of the Census.

HOLT, M. (1986).
Evaluation Report on the Training Materials and Methods for the Survey of Income and Program Participation. Forthcoming SIPP Working Paper.

HUBBLE, D. and D. JUDKINS (1986).
Preliminary Quality Assessment of Gross Flow Estimates from SIPP. Paper to be presented at the 1986 Meeting of the American Statistical Association.

HUGGINS, V. (1986)
A Comparison of Seven Imputation Procedures for ISDP. Forthcoming SIPP Working Paper.

HUGGINS, V., M. SAMUHEL, and L. WEIDMAN (1985).
Methodologies for Imputing Longitudinal Survey Items. SIPP Working Paper Series No. 8506, U.S. Bureau of the Census.

HUGGINS, V. and L. WEIDMAN (1986a).
An Investigation of Model-Based Imputation Procedures Using Data from the Income Survey Development Program. Forthcoming SIPP Working Paper.

HUGGINS, V. and L. WEIDMAN (1986b).
An Investigation of the Imputation of Monthly Earnings for the Survey of Income and Program Participation Using Regression Models. Forthcoming SIPP Working Paper.

JUDKINS, D., D. HUBBLE, R. SINGH, D. MCMILLEN, and L. ERNST (1984).
Weighting of Persons for SIPP Longitudinal Tabulations.
Proceedings of the Survey of Research Methods Section, American Statistical Association, Washington, D.C.

KALTON, G. (1985).
Handling Wave Nonresponse in Longitudinal Surveys.
Proceedings of the U.S. Bureau of the Census First Annual Research Conference.

KALTON, G. and J. LEPKOWSKI (1983).
Cross-Wave Item Imputation. In Technical, Conceptual and Administrative Lessons of the Income Survey Development Program (ISDP), M. David, ed. New York: Social Science Research Council.

KALTON, G., J. LEPKOWSKI, and T. LIN (1985).
Compensating for Wave Nonresponse in the 1979 ISDP Research Panel.
Proceedings of the Survey of Research Methods Section, American Statistical Association, Washington, D.C.

KALTON, G., D. MCMILLEN, and D. KASPRZYK (1986).
A Review of Nonsampling Error Issues in the Survey of Income and Program Participation. Proceedings of the U.S. Bureau of the Census Second Annual Research Conference.

KALTON, G. and M. MILLER (1986).
Effects of Adjustment for Panel Nonresponse On Survey Estimates. Paper be presented at the 1986 Meeting of the American Statistical Association.

KOBILARCIK, E. and R. SINGH (1986).
SIPP Longitudinal Estimation for Persons' Characteristics.
Paper to be presented at the 1986 Meeting of the American Statistical Association.

LAMAS, E.J. and J. MCNEIL (1984).
The Measurement of Household Wealth in SIPP.
Proceedings of the Social Statistics Section, American Statistical Association, Washington, D.C.

MCARTHUR, E. and K. SHORT (1985).
Characteristics of Sample Attrition in the Survey of Income and Program Participation. Proceedings of the Survey of Research Methods Section, American Statistical Association, Washington, D.C.

MCARTHUR, E. and K. SHORT (1986).
Measurement of Attrition from the SIPP through the Fifth Wave of the 1984 Panel. Memorandum to the Distribution List, April 10, 1986, and forthcoming Working Paper.

MCMILLEN, D. and R. HERRIOT (1985).
Toward A Longitudinal Definition of Households.
Journal of Economic and Social Measurement, Vol. 13, No. 3-4.

MCMILLEN, D. and D. KASPRZYK (1985).
Item Nonresponse in the Survey of Income and Program Participation.
Proceedings of the Survey of Research Methods Section, American Statistical Association, Washington, D.C.

MCMILLEN, D., D. KASPRZYK, and G. KALTON (1986)
A Further Examination of Nonsampling Error Issues for Panel Surveys.
Paper to be presented at the 1986 Meeting of the American Statistical Association.

MOORE, J. (1986).
SIPP Record Check Analysis Strategy. U.S. Bureau of the Census Memorandum, January 31, 1986.

MOORE, J. and D. KASPRZYK (1984).
Month-to-Month Recipiency Turnover in the ISDP.
Proceedings of the Survey of Research Methods Section, American Statistical Association, Washington, D.C.

NELSON, D., D. MCMILLEN, and C. BOWIE (1986).
Noninterview Rates in SIPP. Forthcoming SIPP Working Paper.

NELSON, D., D. MCMILLEN, and D. KASPRZYK (1985).
An Overview of the Survey of Income and Program Participation: Update 1.
SIPP Working Paper Series No. 8401, U.S. Bureau of the Census.

RYSCAVAGE, P. and J. BREGGER (1985).
New Household Survey and the CPS: A Look at the Labor Force Differences.
Monthly Labor Review, September 1985.

RYSCAVAGE, P. and K. SHORT (1985).
SIPP Labor Force Transitions: Problems and Promises.
Paper prepared for presentation at the 1985 Meeting of the American Economic Association. Forthcoming SIPP Working Paper.

SAMUHEL, M. and V. HUGGINS (1984).
Longitudinal Item Imputation in a Complex Survey.
Proceedings of the Survey of Research Methods Section, American Statistical Association, Washington, D.C.

SATER, D. (1985).
Enhancing Data from the Survey of Income and Program Participation with Data from Economic Censuses and Surveys.
SIPP Working Paper Series No. 8505, U.S. Bureau of the Census.

SHORT, K. and E. MCARTHUR (1986).
Life Events and Sample Attrition in the Survey of Income and Program Participation.
Paper to be presented at the 1986 Meeting of the American Statistical Association.

SHORT, K. and K. WOODROW (1985).
An Exploration of the Applicability of Hazards Models in Analyzing the Survey of Income and Program Participation: Labor Force Transitions. Proceedings of the Social Statistics Section, American Statistical Association, Washington, D.C.

SMITH, P. (1984).
Variance Estimation Software Packages.
U.S. Bureau of the Census draft report.

ST. CLAIR, J. (1985).
Survey of Income and Program Participation 1984 Panel: Results of the Reinterview Program for Waves 2 through 4. U.S. Bureau of the Census Memorandum, July 2, 1985.

U.S. Bureau of the Census, Current Population Reports, Series P-70, No. 1, Economic Characteristics of Households in the United States; 3rd Quarter 1983, U.S. Government Printing Office, Washington, D.C. 1984.

U.S. Bureau of the Census, Current Population Reports, Series P-70, No. 2, Economic Characteristics of Households in the United States; 4th Quarter 1983, U.S. Government Printing Office, Washington, D.C. 1985.

U.S. Bureau of the Census, Current Population Reports, Series P-70, No. 3, Economic Characteristics of Households in the United States; 1st Quarter 1984, U.S. Government Printing Office, Washington, D.C. 1985.

U.S. Bureau of the Census, Current Population Reports, Series P-70, No. 4, Economic Characteristics of Households in the United States; 2nd Quarter 1984, U.S. Government Printing Office, Washington, D.C. 1985.

U.S. Bureau of the Census, Current Population Reports, Series P-70, No. 5, Economic Characteristics of Households in the United States; 3rd Quarter 1984, U.S. Government Printing Office, Washington, D.C. 1985.

U.S. Bureau of the Census, Current Population Reports, Series P-70, No. 6, Economic Characteristics of Households in the United States; 4th Quarter 1985, U.S. Government Printing Office, Washington, D.C. 1986.

VAIANA, M. (1985).
A Product Strategy for SIPP.
Paper prepared for the Census Bureau.

WEIDMAN, L. (1986).
Investigation of Gross Changes in Income Recipiency from the Survey of Income and Program Participation. Paper to be presented at the 1986 Meeting of the American Statistical Association.