Survey of Income and Program Participation and Related Longitudinal Surveys: 1984

Selected Papers given at the
1984 Annual Meeting of the
American Statistical Association
in Philadelphia, Pa.
August 13-16, 1984

Department of Commerce
Bureau of the Census
Population Division
January 1985

SURVEY OF INCOME

AND

PROGRAM PARTICIPATION

AND RELATED LONGITUDINAL SURVEYS: 1984

Selected Papers

given at the

1984

Annual Meeting

of the

American Statistical Association

in Philadelphia, Pa.

August 13-16, 1984

Compiled by:

DANIEL KASPRZYK
DELMA FRANKEL

Population Division
Bureau of the Census
Department of Commerce

January 1985

U.S. BUREAU OF THE CENSUS
John G. Keane, Director
C.L. Kincannon, Deputy Director
William P. Butz, Associate Director for Demographic Fields

POPULATION DIVISION
Roger A. Herriot, Chief

Acknowledgments

This publication is composed of papers prepared by many different authors for presentation at the American Statistical Association meetings, August 13-16, 1984. We would like to thank these authors for their cooperation in making the papers available for publication. (The views expressed are the authors' and do not necessarily represent the views of the Bureau of the Census.) These papers have been compiled and prepared for publication by Daniel Kasprzyk and Delma Frankel, Population Division. Clerical and editorial assistance was provided by Hazel Beaton.

Suggested Citation

"Survey of Income and Program Participation and Related Longitudinal Surveys: 1984," compiled by Daniel Kasprzyk and Delma Frankel. U.S. Bureau of the Census, Washington, D.C. 1984.

PREFACE

This report is comprised of a selection from the papers presented at the 144th Annual Meeting of the American Statistical Association (ASA) in Philadelphia, Pennsylvania, August 13-16, 1984. Most of the papers in the volume deal with methodological and substantive studies related to the Survey of Income and Program Participation (SIPP) and the Income Survey Development Program (ISDP), an experimental program designed to test procedures used in conducting SIPP.

The SIPP is a new Census Bureau survey collecting data that will help measure income distribution and poverty throughout the country more accurately. These data will be used to study Federal and State aid programs (such as food stamps, Social Security, Supplemental Security Income, Aid to Families with Dependent Children, Medicaid, Medicare, and others), to estimate future program costs and coverage, and to assess the effects of proposed changes in program eligibility rules or benefit levels.

Households in the survey will be interviewed at 4-month intervals over a period of 2 1/2 years. The reference period will be the 4 months preceding the interview. In all, about 20,000 households will be interviewed, approximately 5,000 each month. Field operations will be handled through the Census Bureau's 12 regional offices.

Recurring questions will deal with employment, types of income, noncash benefits, assets, liabilities, and taxes. Periodic questions will be added dealing with school enrollment, marital history, migration, disability, and other topics. Additional supplemental questions will also be added to the SIPP questionnaire as the need arises.

Preliminary tabular data from the SIPP is distributed as part of the Current Population Report series P-70, Economic Characteristics of Households in the United States. The first report covering the third quarter of 1983 was released in September 1984. In addition, the first in a series of microdata files from the SIPP is now available. This file contains the results of interviews conducted between October 1983 and January 1984. Microdata files for Waves 2 through 9 of the 1984 SIPP Panel will be released approximately every 4 months, with Wave 2 appearing in February 1985.

Contents. As shown in the Table of Contents, the grouping of the papers (and accompanying discussion comments) basically is in keeping with the ASA sessions at which the presentations were originally given. There were four "Survey of Income and Program Participation" sessions, two of which were included in the Social Statistics Section and two in the Survey Research Methods Section of the ASA meetings. These sessions covered a range of topics, both methodological and substantive, about longitudinal surveys and SIPP. Finally, the Social Statistics Section sponsored a session on "Case Studies in Panel Survey Design: The International Experience." This session provided an opportunity for individuals involved in the design and development of longitudinal surveys to share experiences and discuss issues of mutual interest.

The contents of the report were prepared by the individual authors for publication in the 1984 American Statistical Association Proceedings. For this reason, the format conforms basically to that required by ASA.

Acknowledgements. This publication is composed of papers prepared by many different authors for presentation at the ASA meetings, August 13-16, 1984. We would like to thank these authors for their cooperation in making the papers available for publication. Clerical and editorial assistance was provided by Hazel Beaton.

SURVEY OF INCOME AND PROGRAM PARTICIPATION AND RELATED LONGITUDINAL SURVEYS: 1984

Survey of Income and Program Participation: Session I

Chair: PAUL PLANCHON
Mathematica Policy Research, Inc.

Discussion

Survey of Income and Program Participation: Session II

Chair: GEORGE HALL
Baseline Data Corporation

Survey of Income and Program Participation: Session III

Chair: DANIEL HORVITZ
Research Triangle Institute

Survey of Income and Program Participation: Session IV

Chair: GRAHAM KALTON
Survey Research Center
University of Michigan

Discussion

Case Studies in Panel Survey Design: The International Experience

Session V

Chair: BARBARA A. BAILAR
Associate Director for Statistical Standards and Methodology
Bureau of the Census

SURVEY OF INCOME

AND

PROGRAM PARTICIPATION:

SESSION I

This section is comprised of five papers presented in this session which was sponsored by the Section on Social Statistics.

AN ANALYSIS OF INTRA-YEAR UNEARNED INCOME FLOWS ON THE ISDP

PAT DOYLE, MATHEMATICA POLICY RESEARCH, INC.

This paper presents excerpts from a study conducted for the Food and Nutrition Service, USDA (Doyle, 1984). The overall objective of the study was to provide the foundation upon which improvements could be made to the simulation of Food Stamp Program costs and caseloads. The simulations are currently conducted on CPS-type files which lack detail on intra-year income receipts. The Food Stamp Program, however, has a monthly accounting period, thus requiring that the simulation model first allocate annual income to monthly amounts. In so doing, the current food stamp model relies on simple assumptions concerning variation in unearned income. These assumptions were made in the absence of information on intra-year income flows. Essentially the model assumes that with the exception of welfare and unemployment compensation benefits, unearned income is received evenly throughout the year. Examination of data from the ISDP shows, as expected, that for some sources this assumption is sufficient. However, for other income sources, the observed intra-year income fluctuations are large. Therefore, in order to improve the model of the Food Stamp Program, existing procedures for allocation of unearned income should be changed. This presentation reports on the monthly variation in income recipiency and amounts for unemployment compensation, workmen's compensation, asset income, and other unearned income exclusive of welfare.

THE DATA

The focus of this research was guided by the information available on CPS-type files. One important source of information on these files which contributes to the understanding of intra-year patterns of recipiency is the duration of the work period within a calendar year. That is, part-year workers can be distinguished from full-year workers. Hence, one dimension of the descriptive statistics which follow is work versus non-work.

Furthermore, since the improvements to the simulation do not extend to the imputation of fluctuations in household composition within the year on the CPS, an analysis file was designed to replicate the cross-sectional nature of the CPS files. That is, individuals in households as they existed in Wave 5 were extracted along with retrospective longitudinal data on income receipts during the calendar months of 1979. In order to target the research to the potentially eligible population, the universe for the study was further restricted to low income households. Low income households were defined as households with composition defined as of the Wave 5 interview where the total monthly income aggregated over all members was less than twice the Food Stamp eligibility limit for at least one month during calendar year 1979.[1] For the initial screen, if an observation reported recipiency for an income source but did not report the dollar amount it was counted as receipt of a zero amount. Individuals not present in the sample for the entire reference period were treated as non-recipients of income during their absence. There were 9,383 individuals extracted, 7,468 of which were present in the ISDP sample all 12 months of calendar year 1979.

THE RESULTS

Workmen's Compensation. For this source the ISDP provided mixed results. There were extremely few observations (71 individuals reporting receipt during the year, 65 of whom were part-year workers). Furthermore, individual reported amounts were observed to be unreasonable (one observation reported receipt of $16,666 in each of three months). The question still remains, therefore, as to whether or not to allocate this income source solely to periods of non-work. The ISDP data showed that more than half of the respondents (53% based on weighted counts) reported receipt during the work period. Of those, 57% received workmen's compensation only during the work period, the remainder received it during both work and non-work periods.

For purposes of modeling food stamp eligibility, this income source was lumped with the residual other income category discussed below. As described there, duration of receipt was first randomly imputed and then the actual calendar months of recipiency were randomly assigned. This stochastic process was independent of the assignment of work versus non-work periods.

Unemployment Compensation. In the analysis of unemployment compensation the issue of whether or not to allocate recipiency evenly over the year is irrelevant because this source is directly related to labor force activity. It is currently assumed that when both recipients of unemployment compensation and weeks of unemployment are reported, that the periods coincide. Therefore, the key issue is to what extent does UI recipiency coincide with work periods. Table 14 displays recipients of unemployment compensation by duration of the work period and by the coincidence of work and receipt of this income source. Months working in this case (as in all other cases in this report) are months during which earnings are reported. This table shows that 17.6% of UI recipients in poor households were full-year workers. An additional 42% were part-year workers reporting UI receipt while working. This seems unusually high compared to other estimates. Burtless (1983) reports that 6% of UI recipients in a given week are underemployed rather than unemployed. It is useful, therefore, to examine the recipients more closely. Table 15 shows the duration of the coincidence of work and UI receipt for the 2.6 million workers. In over 95% of the cases the periods of receipt of UI did not encompass the full period of work. Over half of this group only received unemployment compensation during one of the months reported working. An additional 30% reported receipt only during two months of working. This suggests that the coincidental period was one of transition. That is, either a job was lost early in the month and UI benefits are received later in the month or UI was received early in the month and a job was secured later. The 54% of the cases with one month of overlap reflect individuals with only one transition period and the remaining 30% represent cases with two transition periods. Fifteen percent of the cases had more than two months of coincidental receipt of earnings and UI. These include both cases where more than two transitions occurred as well as underemployed individuals.

These results suggest that a significant number of the cases reporting earnings and UI receipt concurrently were in a transition period. This is further supported by the fluctuation in average benefits displayed in Table 16. average UI benefits received during the worked period were $190 whereas average benefits received during the nonwork period were $316. When disaggregated by whether or not individuals received UI benefits during both periods of work and nonwork, the contrast in average benefits across the two periods is increased. Recipients during both periods report an average of $191 in the work period and $331 in the nonwork period. Persons receiving benefits solely during non-work periods reported an average of $304 and persons receiving benefits solely during work periods reported $189.

Based on these results the allocation of UI within the year on CPS-type files will be modified. The process will be to first derive expected duration of receipt of UI from the reported annual amount incorporating regulations regarding maximum weekly benefits and maximum weeks of receipt.[2] Then the period of receipt will be allocated to periods of unemployment. If the duration of receipt exceeds the unemployment period, excess months of recipiency will be allocated to periods of work immediately surrounding periods of unemployment.

Asset Income. Table 17 shows the frequency counts of persons in low income households by number of months receiving income from assets. This category includes interest income from savings accounts, etc., dividend income from stocks and bonds, and rental income from property. Overall, 55% of the population reporting asset income received that income the full year. Of the remaining persons with asset income, the distribution by number of months received is biased due to the way in which asset income data were gathered in the ISDP. For many sources included here, lump sum quarterly or semi-annual amounts were requested by the interviewer depending on whether the individual fell in the three month recall sample or the six month recall sample. Recipiency for these cases was evenly allocated over the relevant period thus introducing a bias into the examination of recipiency patterns within the year.

Table 17 also shows the differences in recipiency patterns for elderly and non-elderly (elderly in this case refers to persons age 60 and older). For the older population, the proportion of individuals receiving asset income for 12 months is 72%, which is significantly higher than the overall population. Forty-nine percent of the non-elderly recipients received asset income the full year.

Based on the fact that such a high proportion of the asset income recipients received that source the full year and that the ISDP questionnaire design prevents a detailed analysis of intra-year patterns of recipiency, asset income recipiency on the FNS data base will be evenly allocated across calendar months.

Table 18 shows the change in average monthly asset income between work and non-work periods for persons in low-income households who received asset income during both periods and for persons whose assets were less than $3,000. The asset screen was imposed to further isolate individuals who were potentially eligible to receive food stamps. For the potentially eligible group, asset income did not vary significantly across periods of work and non-work. Potentially eligible elderly individuals reported $10 on average during both periods while the non-elderly reported $5 during the nonwork period and $4 during the work period.

Based on the observation that individuals in low income households with small amounts of assets tend not to have significant variation within the year, asset income amounts in the FNS data base will be evenly allocated across calendar months.

Other Income. The pattern of recipiency of other income varies significantly for elderly and non-elderly recipients. Table 19 shows that 77% of the elderly report receiving other income continuously throughout the year whereas only 24% of non-elderly received this income amount for the full 12 months. Recipiency patterns for non-elderly were further disaggregated by labor force activity but then did not appear to be a significant difference between patterns of recipiency of part-year workers and the rest of the non-elderly population.

Recognizing that this residual income category includes income sources such as lump sum payments which are received intermittently throughout the year as well as regularly received sources such as pensions and social security, the non-elderly recipients were further examined by type of other income received. Table 21 shows the distribution of non-elderly recipients in poor households by number of months receiving regular and irregular income. Regular income consists of social security, railroad retirement, veterans benefits, and other disability payments excluding workmen's compensation, and private and government pensions. Irregular income includes all other sources not classified in the previous sections or classified as regular income. Over one-third of the recipients of regular income received that source for 12 months whereas only 15 percent of the recipients of irregular income received that source for the full year. In both cases, the majority of the remaining recipients fall at the lower end of the distribution. The balance of the recipients in both cases are probably fairly evenly distributed but apparent reporting problems on the ISDP tend to bias the observed distribution. The apparent reporting problems result from respondents' tendency to report changes in recipiency status more often between waves than between consecutive pairs of months (Moore and Kasprzyk, 1984).

Table 22 shows the change in average amounts received for other income across periods of work and non-work separately for elderly and non-elderly recipients. Overall, average income amounts were 20% higher during work periods. As was true for recipiency, there was a significant difference between elderly and non-elderly income flows. For the elderly in low income households who received other income during work and non-work periods, amounts received while working were approximately the same as those received while not working. When restricted further to elderly individuals with low assets, more variation existed but average amounts received while working were only six percent less than amounts received while not working. Non-elderly individuals, on the other hand, reported average benefits received while working to be 130 percent of those received while not working. When restricted to individuals with low assets, the variation was reduced but still remained high with amounts received during work periods 20 percent higher than amounts received while not working.

Again this analysis of the ISDP data suggest that for the elderly an even allocation of other income across months would be sufficient for modeling food stamp eligibility. However, for non-elderly, this method would understate the variation in amounts received from this source.

Based on the distribution presented here, the allocation of other income recipiency on the Food Stamp data bases will be a stochastic process whereby duration is randomly determined and then a period of recipiency is stochastically assigned within the twelve month reference period.[3] The probability upon which the duration will be determined will be derived from the distributions presented in Tables 19 through 21. Recognizing that the irregularity of the distributions is caused, at least to some extent, by reporting patterns on the ISDP, they will be smoothed out in the construction of a cumulative probability function. Also, workmen's compensation will be included with regular income. Amounts received will be evenly allocated over the imputed period of recipiency.

FOOTNOTES

[1]The food stamp eligibility limits were those in effect July 1, 1979. These reflect the OMB poverty guidelines for mid 1978 updated by the CPI.

[2]As is true for other unearned income sources, the CPS does not measure duration of receipt.

[3]As is true for all unearned income sources, the CPS does not measure duration of receipt of other income.

REFERENCES

Burtless, Gary. "Why is Insured Unemployment So Low?" In Brookings Papers on Economic Activity. (Washington, D.C.: The Brookings Institute, 1983)

Doyle, Pat. An Analysis of Intra-Year Income Flows. Final Report to USDA. (Washington, D.C.: Mathematica Policy Research, July 1984)

Moore, J. C. and Kasprzyk, D. "Month to Month Income Recipiency Changes in the ISDP." Paper prepared for the American Statistical Association Meetings, August 1984.

TABLE 14

RECIPIENTS OF UNEMPLOYMENT COMPENSATION

Work Status	Receipt Period	Recipients Count (1000)	Recipients %
Full Year Workers	Work only	778.27	17.6%
Nonworkers	Nonwork only	198.31	4.5
Part Year Workers	Both	1679.63	37.9
	Nonwork only	1595.64	36.0
	Work only	182.46	4.1
Total		4434.31	100.0

SOURCE: Prepared by Mathematica Policy Research using the ISDP/RAMIS II system.

NOTE: This table is based on 371 observations who reported unemployment compensation and who were present in the sample the full year.

TABLE 15

RECIPIENTS OF UNEMPLOYMENT COMPENSATION WHILE WORKING

Number of Months Receiving UI While Working	Number of Months Worked	Recipients Count (1000)	Recipients %
1	1	34.6	1.3%
	2	91.3	3.5
	3-6	126.5	4.8
	7-11	846.8	32.1
	12	338.4	12.8
	Total	1,437.6	54.4
2	2	0	0
	3-6	144.3	5.5
	7-11	367.4	13.9
	12	283.2	10.7
	Total	794.9	30.1
3+	3-6	36.1	1.4
	7-11	215.1	8.1
	12	156.6	5.9
	Total	406.9	15.4
TOTAL		2640.4	100.0

SOURCE: Prepared by Mathematica Policy Research using the ISDP/RAMIS II system.

NOTE: This table is based on 210 observations who were interviewed over the entire calendar year. There were an additional 27 observations reporting receipt of UI while working who were not interviewed during at least one wave.

TABLE 16

AVERAGE UNEMPLOYMENT COMPENSATION FOR PART-YEAR WORKERS BY WORK STATUS

Period of Receipt	Months Worked	Months Not Worked
Work Only	189	0
Nonwork only	0	304
Both	191	331
TOTAL	191	316

SOURCE: Prepared by Mathematica Policy Research using the ISDP/RAMIS II system.

NOTE: This table is based on 334 observations reporting UI receipt who also reported amounts, that is, cases of nonresponse have been screened out.

TABLE 17

DURATION OF RECEIPT OF ASSET INCOME

Number of Months	Elderly Count (1000)	Elderly %	Nonelderly Count (1000)	Nonelderly %	Total Count (1000)	Total %
1	176.1	1%	1096.7	2%	1272.8	2%
2	338.6	2	746.3	2	1084.9	2
3	290.0	2	2305.6	5	2595.6	4
4	308.1	2	1060.2	2	1368.3	2
5	264.8	1	968.4	2	1233.2	2
6	961.7	5	5991.9	13	6953.6	11
7	648.6	4	1831.6	4	2480.2	4
8	371.7	2	2597.4	6	2969.1	5
9	848.8	5	4034.3	9	4883.1	8
10	639.0	3	1191.9	3	1830.9	3
11	237.0	1	1309.9	3	1546.9	2
12	13187.7	72	21902.4	49	35090.1	55
Total	18272.1	100	45036.6	100	63308.7	100

SOURCE: Prepared by Mathematica Policy Research using the ISDP/RAMIS II system.

NOTE: This table is based on 5992 observations, 1664 of which are elderly. Persons not present in the sample the full year have not been screened.

TABLE 18

CHANGE IN AVERAGE ASSET INCOME FOR RECIPIENTS OF ASSET INCOME WHILE WORKING AND NOT WORKING

	Average Monthly Asset Income			
			Assets < $3000	
Age	Months Working	Months Not Working	Months Working	Months Not Working
Elderly	125	70	10	10
Not Elderly	22	27	5	4
Total	33	32	6	5

	Unweighted Number of Persons upon which Averages are Based			
			Assets < $3000	
Age	Months Working	Months Not Working	Months Working	Months Not Working
Elderly	138	107	75	55
Not Elderly	1,158	929	939	753
Total	1,296	1,036	1,014	808

SOURCE: Prepared by Mathematica Policy Research using the ISDP/RAMIS II system.

NOTE: This table consists of part-year workers who reported receiving asset income in both periods of work and periods of nonwork. Persons not interviewed the full year have not been screened.

TABLE 21

DISTRIBUTION OF NON-ELDERLY RECIPIENTS OF OTHER INCOME BY DURATION OF RECEIPT BY TYPE OF OTHER INCOME RECEIVED

Number of Months Receiving	Irregular		Regular	
	Count (1000)	%	Count (1000)	%
1	4931.5	32	954.3	13
2	1825.1	12	547.3	7
3	2316.3	15	788.1	11
4	765.1	5	177.9	2
5	1040.9	7	327.0	4
6	384.1	2	524.2	7
7	456.8	3	122.9	2
8	259.0	2	161.4	2
9	835.0	5	476.4	6
10	129.3	1	156.6	2
11	355.8	2	311.7	4
12	2346.1	15	2801.5	38
TOTAL	15645.0	100	7349.2	100

SOURCE: Prepared by Mathematica Policy Research using the ISDP/RAMIS II system.

TABLE 22

CHANGE IN AVERAGE MONTHLY OTHER INCOME ACROSS PERIODS OF WORK AND NONWORK

	Average Monthly Other Income			
			Assets < $3000	
Age	Months Working	Months Not Working	Months Working	Months Not Working
Elderly	330	325	302	321
Not Elderly	452	348	418	346
Total	407	339	381	338

	Unweighted Number of Persons upon which Averages are Based			
			Assets < $3000	
Age	Months Working	Months Not Working	Months Working	Months Not Working
Elderly	204	201	137	133
Not Elderly	349	332	306	291
Total	553	533	443	424

SOURCE: Prepared by Mathematica Policy Research using the ISDP/RAMIS II system.

AN ANALYSIS OF TURNOVER IN THE FOOD STAMP PROGRAM

Irene Smith Lubitz, Mathematica Policy Research
Timothy J. Carr, U.S. General Accounting Office

The findings reported in this paper are based on research carried out by Tim Carr, Pat Doyle and Irene Smith Lubitz as a part of MPR's analysis of participation in the Food Stamp Program, pursuant to contract No. 53-3198-0-101 with the Food and Nutrition Service, USDA. The authors are indebted to Steven Carlson, Harold Beebout, Jim Ohls and others for helpful comments on the research on which this paper is based. The conclusions presented here are solely those of the authors and do not necessarily reflect the opinions of the Food and Nutrition Service or Mathematica Policy Research.

INTRODUCTION

Researchers and policymakers have long had an interest in analyzing the factors associated with the decision of eligible individuals and households to participate (or not to participate) in income maintenance programs. Accordingly, a large and growing body of research on this topic has appeared in recent years. There has also been considerable interest in patterns of program participation over time (i.e., entry into, and exit from, these programs). However, there have been relatively few studies of the longitudinal patterns of participation (which we hereafter refer to as "turnover"), because of stringent data requirements. Such studies that have appeared have usually been based on peculiar samples (e.g., participation in negative income tax experiments).

In this paper we report the findings of an analysis of turnover in the Food Stamp Program based on data from the 1979 Income Survey Development Program Research Panel (hereafter referred to as the ISDP data). The ISDP data base was a precursor of the forthcoming Survey of Income and Program Participation, and displays many of the same attributes that will make SIPP especially advantageous for research on turnover. As such, our research illustrates the potential uses of SIPP for research in this area.

THE DATA

A successful analysis of food stamp turnover requires data possessing several essential characteristics. To our knowledge, the ISDP is the first data base to exhibit all of these desirable features simultaneously. These characteristics are as follows:

- o The ISDP is nationally representative, and it yields enough observations (about 7,500 households) to permit meaningful analysis.
- o The ISDP is longitudinal; retrospective interviews were conducted quarterly for a period of fifteen months.
- o Data on income that a sample household receives from various sources (including the receipt of food stamps) is ascertained on a month-by-month basis. Furthermore, the exact timing of all changes in household composition between interviews is pinpointed. These data requirements are especially crucial to the analysis of turnover in the Food Stamp Program because many households enter and exit the program in the space of less than a year, and because program participation may be triggered by events such as a temporary drop in household income that could not be picked up by a survey that only measures a household's annual income.
- o The data on income, household composition, and assets permit us to simulate eligibility for the program, using the eligibility criteria actually used in 1979.
- o There is a wealth of explanatory variables, such as household and individual characteristics, types of income, and program participation, available to the analyst for both tabular and multivariate analyses of turnover.

METHODOLOGY

We focus primarily on three measures of program turnover:

- o The <u>entry rate</u>; that is, the probability that a household that was not receiving food stamps in a given month received food stamps the next month.
- o The <u>exit rate</u>; that is, the probability that a household that received food stamps in a given month did not receive food stamps in the next month.
- o The <u>annual/monthly ratio</u>; that is, the ratio of the probability that a household receives food stamps over the course of a year to the probability that it receives food stamps in a given month.

Both tabular and multivariate techniques were used in the analysis. The multivariate analysis used the RATE model for the analysis of event histories that has been developed by Nancy Tuma and her colleagues (Tuma et al., 1979), and applied in studies of marital instability, unemployment duration, and other socioeconomic phenomena. Although the tabular analysis of exit and entrance rates focused on one characteristic at a time, making it difficult to sort out compositional effects, the multivariate analysis in general yielded largely similar results.

EMPIRICAL FINDINGS

Tabular Analysis

The evidence from the ISDP panel, via tables and multivariate analysis, is that turnover in food stamp participation is high. We estimate the ratio of annual to monthly participation at 1.74, indicating that the number of households served by the program over the course of a year is about 70 percent greater than the caseload in an average month (Table 1). To illustrate the implications of this finding, we note that program data for 1979 indicate that the average monthly caseload in 1979 was 6.5 million households (USDA, 1979). This annual-monthly ratio implies that 11.3 million households--about 14 percent of all households--received food stamps sometime during 1979.

Most of the food stamp households observed in our data received food stamps for only part of the year. About two thirds of the sample households who received food stamps during 1979 participated in the program for less than the full year and nearly a third of the participants received food stamps for 3 months or less in 1979.[1] Only about one-third of all food stamp recipient households observed received food stamps "continuously" (that is, for all months present in the sample). In other words, a truly "long-term caseload" may account for only about a third of the households who receive food stamps in a given year.

The 1979 sample period is rather short for observing individual households' food stamp spells over time. However, even during this relatively short observation period, about 12 percent of food stamp households experienced more than one spell[2] of participation. This would seem to indicate that recidivism in the Food Stamp Program--households returning to the program after not participating for some interval--may be high.

The average monthly rate of exit from the Food Stamp Program is estimated at over seven percent. That is, in a given month, over seven percent of the caseload may be expected to leave the program by the next month. The exit rate in a given month is the proportion of the previous month's caseload (or of a caseload subset) that has now left the program, and is estimated from caseload counts and exits from the program[3] in each month of 1979. In the aggregate, then, a substantial share of the caseload "turns over" each month, with perhaps 500 thousand households[4] leaving the program and being replaced, in a steady state of no program growth, by a similar number of new entrants. (When the program is expanding, entrances will exceed exits, and if contracting exits will exceed entrances.) In fact, there was significant expansion in 1979, due to programmatic changes.

The program entrance rates measure inflows to food stamp participation. These rates, expressed relative to total population, are much lower than exit rates but in fact represent flows into the program that approximately equal outflows (exits).[5] The average monthly entrance rate as shown in Table 1 is 0.5% per month--that is, the average probability of a nonparticipant in a given month becoming a participant in the next month is about half of one percent.

Turnover rates in the Food Stamp Program, however measured, appear to be quite different for different kinds of households. The various measures of turnover presented for different population subgroups indicate that the more "permanent" part of the food stamp caseload includes households participating in AFDC and other welfare programs, and elderly households. A more transient group of participants includes younger non-welfare households with more labor force attachment and education.

Multivariate Analysis

The multivariate analysis of turnover in program participation, using the RATE model, provides estimates of the independent association of household characteristics with different turnover rates. The results of estimating our basic model of transitions to and from participation in the Food Stamp Program are presented in Table 2. The precise interpretation of these coefficients is not entirely straightforward, as entry and exit rates are complicated functions of the coefficients.[6] Note that the qualitative effect of an explanatory variable on entry and exit rates is indicated by the sign of its coefficient, just as would be the case in the more familiar linear regression model. For instance, the coefficient of the elderly/disabled dummy variable is positive in the entry model, and negative in the exit model. This indicates that households containing elderly or disabled persons are more likely to enter the program, and less likely to exit from it, *ceteris paribus*.

In general, the results are consistent with the results of the tabular analysis presented above, in that the household characteristics that appear to be associated with high entry and exit rates on the basis of the tabular analysis are also those that appear to be associated with high entry and exit rates on the basis of the multivariate analysis. In particular, the following findings are both statistically and substantively significant.

- Nonwhite households who are not in the program are far more likely to enter the program in any given month than otherwise similar white households; furthermore, nonwhite households that are receiving food stamps in any given month are likely to stay on the program longer (i.e., have lower exit rates) than otherwise similar white households.
- Households within which there is no currently employed person are both more likely to enter the program and less likely to exit the program, *ceteris paribus*.
- Households with one head and households with an elderly or disabled person tend to stay on the program longer than other households, all other things being equal.
- Households that receive AFDC are both more likely to enter the program, and less likely to leave, than otherwise similar households.

This last finding, especially the higher entry rate of AFDC households, is especially interesting because it has been hypothesized by some previous researchers that there is a "stigma" effect that acts as a sort of psychological barrier to participation in income maintenance programs (e.g., see Czajka, 1981). These researchers have found that participation in one program is generally correlated with participation in other programs, and our findings tend to confirm theirs. This behavior can be explained in two ways. First, it may be the case that there are households whose members are psychologically less averse to receiving welfare than others, and hence are more likely to apply for benefits from all programs. Second, a household may perceive little or no additional stigma from applying for and receiving benefits from other programs. Of course, these explanations are not mutually exclusive, and it is difficult, if not impossible, to disentangle them with the data available to us.

The estimated coefficients of the RATE model can be used to predict monthly entry and exit rates, annual participation rates, the ratio of annual to monthly participation rates, and the expected duration of spells of participation for a hypothetical household with any combination of characteristics.[7] In order to make the implications of our estimated models more transparent, we have calculated the values of these functions for certain combinations of characteristics.

Specifically, our approach to this presentation is as follows. First, we define a "baseline" household that has characteristics that are fairly typical: a white household with two heads, at least one of whom is employed, and no children. Furthermore, this hypothetical household does not receive AFDC, nor does it contain an elderly or disabled person. We have calculated predicted monthly entry and exit rates and other measures of turnover for the baseline household; these results are presented in the first row of Table 3.

The numbers in the other rows of Table 3 are derived by altering the assumed values of the explanatory variables one by one. For instance, the row labeled "elderly/disabled" pertains to a hypothetical household that contains an elderly or disabled person, but is otherwise similar to the baseline household defined alone, and so forth. As one would expect based on the results in Table 1, there are certain identifiable types of "low-turnover" households, such as households with an elderly or disabled person and households with no person who is currently employed, that are characterized by a low ratio of annual to monthly participation rates and a high predicted duration on the program.

The last three rows illustrate the effect on a household of having two or more characteristics that are associated with low turnover. The first two of these rows simulate the case of a household headed by a single person and containing a child who is under 6 (in the first case) and over 6 (in the second case). The last row describes a hypothetical household consisting of a retired elderly person who lives alone. Our results imply that if he/she receives food stamps, he will be on the program for an average of over four years before exiting, several times longer than the expected duration of participation for the population as a whole.

FOOTNOTES

[1]These estimates are based on households present in sample for the full calendar year. However, when households present for only part of the year are included, the results are similar. Note that these estimates, although illustrative of caseload composition in a given year, do not imply estimates of average duration of spell length due to the restricted sample period. Households with fewer than 12 food stamp months in 1979 may be observed during spells that began before or ended after that year.

[2]Estimate based on "true spells"--spells of participation separated by an interval in which the household was present in the sample but not receiving food stamps.

[3]This calculation is based on "true exits" only; where a true exit is, generally speaking, one where the unit remains in the sample but is observed to be no longer receiving food stamps, as opposed to a unit who leaves the sample following some period receiving food stamps.

[4]This estimate is based on an average monthly caseload of 6.5 million households in 1979, as indicated by program data.

[5]The weighted ISDP counts show about 3.7 million entrances and 3.2 million exits over the course of 1979, consistent with the observed increase in the sample caseload for that period.

[6]The manner in which these functions were calculated is described in Carr et al. (1984, Appendix C).

[7]A detailed description of the manner in which these numbers are calculated is provided in Carr et al. (1984, Appendix C). These calculations assume that the conditions underlying the simple Markov model are satisfied; in particular that the explanatory variables account for all or most systematic variation in entry and exit rates.

REFERENCES

Carr, Timothy J.; Doyle, Pat; and Lubitz, Irene Smith. Turnover in Food Stamp Participation: A Preliminary Analysis. Mathematica Policy Research report submitted to the U.S. Food and Nutrition Service, July 1984.

Czajka, John L. Determinants of Participation in the Food Stamp Program: Spring 1979. Mathematica Policy Research, Draft Final Report submitted to the Food and Nutrition Service, U.S. Department of Agriculture, November 5, 1981.

Tuma, Nancy Brandon; Hannan, Michael T.; and Groeneveld, Lyle P. "Dynamic Analysis of Event Histories." American Journal of Sociology 84, no. 4 (January 1979), pp. 820-854.

U.S. Department of Agriculture, Food and Nutrition Service, Accounting and Reporting Division. Food Stamp Program Statistical Summary of Operations. January-December 1979.

TABLE 1

INDICATORS OF TURNOVER IN FOOD STAMP PROGRAM
PARTICIPATION BY SELECTED UNIT CHARACTERISTICS

	Annual/Monthly Participation Ratio	Exit Rate	Entrance Rate	Spells of Participation		Months of Participation			
				One	Two or More	1-3	4-6	7-11	12
All Households	1.74	7.3%	0.53%	88.5%	11.5%	27.5%	13.8%	25.4%	33.4%
AFDC Status									
Recipient	1.07	2.5	1.13	91.4	8.6	15.8	20.1	34.9	29.2
Nonrecipient	2.01	8.9	0.47	88.0	12.0	29.5	12.7	23.7	34.1
Other Welfare									
Recipient	1.34	3.2	0.86			31.9	10.4	19.3	38.4
Nonrecipient	1.81	7.7	0.51			27.2	14.0	25.8	33.0
Age of Head									
Under 25	1.89	7.1	0.64	91.4	8.6	20.1	12.2	32.3	35.4
25-44	1.81	7.7	0.59	87.5	12.5	26.9	16.0	24.2	32.9
45-59	1.71	7.9	0.41	87.0	13.0	33.5	18.1	18.6	29.9
60-64	1.81	7.7	0.20	88.2	11.8	28.9	8.1	22.6	40.0
65+	1.51	5.7	0.56	90.5	9.5	26.5	6.6	31.2	35.6
Family Status									
Married w/children	1.92	11.7	0.47	84.9	15.1	34.9	24.8	19.3	21.1
Single w/children	1.63	4.7	2.51	88.9	11.1	21.5	11.2	26.9	40.3
Married, no children	1.86	8.5	0.16	92.0	8.0	32.7	10.4	26.9	30.2
Single, no children	1.69	6.4	0.52	90.5	9.5	25.1	6.1	29.6	39.4
Race									
White	1.85	8.2	0.42	88.9	11.1	33.4	13.6	21.8	31.2
Nonwhite	1.56	5.9	1.45	87.8	12.2	16.3	14.2	32.1	37.5
Household Size									
1	1.59	5.6	0.48	89.9	10.1	19.3	3.9	37.2	39.6
2	1.90	7.4	0.32	91.6	8.4	23.0	19.4	17.3	36.2
3-4	1.91	9.3	0.48	87.3	12.7	40.2	13.8	19.4	26.6
5+	1.59	6.1	1.25	86.5	13.5	18.2	16.9	29.4	34.9
Children Under 19									
0	1.73	6.9	0.34	90.8	9.2	27.0	7.2	28.8	37.0
1	1.93	7.4	0.59	86.4	13.6	25.8	25.2	21.1	27.9
2+	1.68	7.6	0.92	87.3	12.7	28.6	14.3	24.3	32.8
Children Under 6									
0	1.77	7.5	0.42	89.1	10.9	27.4	12.9	24.2	35.5
1	1.63	6.3	1.02	85.2	14.8	21.8	17.0	34.1	27.1
2+	1.84	8.1	1.05	92.4	7.6	43.1	11.9	11.1	33.8
Highest Grade Completed									
Less than 9th	1.66	6.0	1.21	89.9	10.1	24.2	9.3	26.2	40.3
9th-11th	1.71	6.1	.90	89.2	10.8	28.6	11.2	21.2	38.9
12th	1.71	9.0	.30	85.7	14.3	22.0	21.1	31.7	25.2
Some college	2.49	13.0	.13	89.4	10.6	51.8	17.9	16.4	13.7
Presence of Earners									
Present	2.00	10.9	.40	87.8	12.2	35.3	15.7	25.3	23.6
Not Present	1.57	4.3	1.08	89.3	10.7	18.7	11.6	25.4	44.3
Elderly or Disabled Persons									
Elderly	1.51	5.3	.43	93.1	6.9	25.5	5.5	30.4	38.5
Disabled	1.86	6.3	3.22	88.3	11.7	23.6	13.6	12.3	28.4
Both	1.46	5.5	.68	81.7	18.3	25.8	8.4	22.9	43.3
Neither	1.82	8.5	.49	87.3	12.7	27.5	16.4	26.6	29.4

Source: Calculated by Mathematica Policy Research from 1979 ISDP Panel. (See text for details of particular calculations.)

TABLE 2

ESTIMATED COEFFICIENTS OF A MODEL OF TURNOVER IN FOOD STAMP PARTICIPATION

Independent Variable	Entry Model		Exit Model	
Constant	-5.374		-2.841	
Elderly/Disabled	.132	(0.92)	-.683	(-3.59)***
Nonwhite	1.601	(14.89)***	-.357	(-2.42)***
Single head	.212	(1.51)*	-.438	(-1.98)**
Youngest child under 6	.793	(4.37)***	-.067	(-0.27)
Youngest child 6-18	.378	(2.13)**	-.037	(-0.14)
AFDC recipient	1.223	(4.26)***	-.349	(-1.62)*
Earner present	-1.353	(-9.71)***	.901	(5.59)***
Single head, child present	.743	(3.48)***	-.333	(-1.14)
x^2	454.24***		116.79***	
Number of observations	7,276		667	

Source: Calculated by Mathematica Policy Research from 1979 ISDP Panel.

Note: Asymptotic t statistics are in parentheses.

* Significant at .10 level (one-tailed test).
** Significant at .05 level (one-tailed test).
*** Significant at .01 level (one-tailed test).

TABLE 3

PREDICTED MEASURES OF TURNOVER IN PARTICIPATION

Household Type	P(Entry)	P(Exit)	Monthly Participation Rate	Annual Participation Rate	Annual/ Monthly Ratio	Predicted Duration
Baseline	0.11	13.4	0.9	2.1	2.46	7.5
Elderly/disabled	0.13	7.0	1.8	3.3	1.76	14.3
Nonwhite	0.56	9.5	5.6	11.3	2.02	10.5
Single head	0.14	8.9	1.6	3.1	1.97	11.3
Youngest child under 6	0.25	12.6	1.9	4.6	2.36	8.0
Youngest child 6-18	0.16	12.9	1.2	3.0	2.41	7.7
AFDC recipient	0.39	9.6	3.9	7.9	2.04	10.4
No earner present	0.45	5.7	7.4	11.8	1.61	17.7
Single head, child under 6	0.67	6.0	10.0	16.3	1.64	16.6
Single head, child 6-18	0.44	6.2	6.6	11.0	1.67	16.2
Single head, elderly, no earner present	0.65	1.9	25.6	30.7	1.20	53.2

Source: Calculated by Mathematica Policy Research from 1979 ISDP Panel.

THE MEASUREMENT OF HOUSEHOLD WEALTH IN THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

Enrique J. Lamas and John M. McNeil, U.S. Bureau of the Census

I. Introduction

The Survey of Income and Program Participation (SIPP) is designed to obtain current estimates of income, labor force activity, and participation in government transfer programs. Currently available survey data have important limitations which have been well recognized and documented [Ycas, 1982; Ycas and Lininger, 1982; Nelson, McMillen, and Kasprzyk, 1983]. The March supplement to the Current Population Survey (CPS) is presently the major source data on the distribution of income. In the March CPS, household members at the time of the interview are asked to recall their income for the previous calendar year. While the CPS survey does well in estimating wage and salary income, it has serious underreporting problems with respect to property income (such as interest, dividend, and rental income) and several other income types (such as Supplemental Security Income (SSI), and worker's compensation)[1]. The CPS has other limitations: subannual income estimates are not available; annual household income estimates cannot be adjusted for changes in household composition or size during the year, for example, as a result of a death, marriage, or divorce; and, information important for assessing economic well-being of the population and for policy analysis, such as assets, taxes and other characteristics, are not covered.

SIPP is designed to overcome many of these limitations. Income data from a wide variety of sources including wage and salary and government transfer programs are collected on a monthly basis. Changes in household composition are also identified on a monthly basis. In addition to improving measures of income and program participation, SIPP will also obtain detailed data on other important topics including asset and liabilities, and taxes. Asset and liability data are important in determining program eligibility and assessing the economic situation of families. Information on the distribution of household wealth is important and changes in wealth provide data on consumer savings. SIPP will be unique among surveys in providing a recurring series on household wealth. The focus of this paper is to describe the effort of SIPP to provide estimates of household wealth holdings in the United States and to present some preliminary results.

Data to study the composition and distribution of household and personal wealth have come from three sources: estate tax returns, synthetic databases, and surveys [Wolff, 1983]. Each source, however, has its limitations. Estate tax returns consist of records filed with the Internal Revenue Service (IRS) for estate tax assessments. Coverage of the population is a major problem for this data source. Only descendents with substantial wealth holdings ($300,000 or more in 1981 and $500,000 or more in 1982) are required to file estate returns. Under certain assumptions about mortality, an "estate multiplier technique" has been used to estimate wealth for living individuals who would have been required to file estate tax returns [Schwartz, 1983]. However, this technique has limited population coverage, and only provides estimates for top wealth holders.

Synthetic databases, such as the Measurement of Economic and Social Performance file (MESP) [Wolff, 1983] and the 1973 Office of Tax Analysis file [Greenwood, 1983], merge data from several sources in order to get appropriate population and asset coverage. The MESP file consists of a statistical match of the 1970 census 1/1,000 public use sample to IRS tax returns. All asset values were imputed from the IRS tax information for financial assets and from the Consumer Expenditure Survey (CES) for consumer durables. The Office of Tax Analysis file consists of a statistical match of the 1973 CPS to tax records from the 1973 Individual Income Tax Model. The IRS data on dividends and interest are used to estimate the market value of financial assets and the IRS data on property taxes are used to estimate the value of real estate holdings. Net wealth estimates are then imputed using a regression of net wealth on asset values based on IRS estate tax data. Limiting these databases are various assumptions which underlie the matching procedures, the procedure of income capitalization, and the extension of estimates to the whole population [Smith, 1983].

A third source of data are surveys covering household assets and liabilities. Major previous surveys include the Survey of Consumer Finances (SCF), the Survey of Financial Characteristics of Consumers (SFCC), and the Income Survey Development Program (ISDP). The SCF are periodic surveys (the latest conducted in 1983) with sample sizes of 3,500 to 4,000 consumer units. The SFCC, conducted in 1962 and 1963, canvassed 2,557 units. Information on the size and components of wealth was collected as of December 31, 1962 [Projector and Weiss, 1966]. Both the SFCC and the 1983 SCF used IRS records in order to oversample high income households. The ISDP, conducted in 1979 and 1980, collected information on assets and liabilities from approximately 7,000 households. The ISDP was a research panel designed to prepare for SIPP. The ISDP interviewed respondents on a quarterly basis on six occasions. In the fifth interview, a supplement on assets and liabilities was administered which collected information as of December 31, 1979.

Surveys have the advantage of being able to have samples and survey instruments specifically designed to gather the necessary detail and information to estimate wealth holdings. Surveys, however, have suffered from limitations in population and asset coverage.

In the next section of this paper, significant features of the SIPP design with respect to the measurement of wealth are discussed. The final section presents results from the first the first wave of SIPP.

II. SIPP Design Features

SIPP is a panel survey consisting of approximately 20,000 households which are interviewed every four months for a period of 2 1/2 years. As SIPP progresses, new panels will be started every year which will allow cross-sectional analyses based on a total sample of approximately 35,000 households. At each interview, information on income, program participation, and other characteristics for each of the previous four months is obtained for each person. Persons who move during the life of the panel, are followed and interviewed at their new addresses.[2] Questionnaire items include a "core" set of questions which are repeated in each wave of interviewing. These items cover labor force participation, detailed income recipiency, and participation in government programs. For waves 2 through 8, the core items are updated and the questionnaire is expanded with additional questions on items not covered in the core. Detailed questions concerning the amounts of personal and household asset and liabilities are included in wave 4 which is to be conducted in September through December 1984. These items will be updated one year later in wave 7. Asset and liability coverage is comprehensive.

The SIPP design is expected to have a positive effect on the ability to measure wealth. Research has found that a major source of bias in survey estimates of wealth is the nonreporting of asset ownership [Ferber, 1982]. Additionally, there is some nonreporting of asset values. Several features of SIPP are expected to have positive effects on the reporting of ownership and value of assets [Radner and Vaughan, 1984].

First, asset ownership questions precede questions on asset values. Separating ownership and amount questions helps focus on identification of asset holdings. In addition, the the relatively sensitive amount questions do not negatively impact on the reporting of asset ownership.

Second, the longitudinal nature of SIPP helps identify asset ownership. Asset ownership information is collected in each wave. In the initial interview, a set of detailed questions designed to identify ownership of income earning assets are asked for each person in the household. An asset roster is created and recorded in the control card. In subsequent interviews, the respondent's asset roster for the previous wave is pretranscribed to the current questionnaire. During an interview, the pretranscribed asset roster is checked for accuracy. Then questions are asked to determine whether any assets have been liquidated or whether any new ones have been acquired. With this procedure, relatively accurate asset ownership information is obtained before respondents are asked about asset values and amounts of liabilities in wave 4.

In longitudinal surveys, attrition, that is dropping out of the sample is of concern. While some respondents leave the sample, there is evidence to suggest that cooperation or rapport obtained in repeated interviews increase the reliability of financial data [Ferber and Frankel, 1978]. Furthermore, the longitudinal nature also provides the opportunity to gather information missed in a previous interview. If one interview is missed for the household or an individual respondent, a "Missing Wave" section is completed. In this section, a limited set of key questions are asked concerning labor force participation, income recipiency, and asset ownership during the missed wave.

Third, ownership, income, and value of assets are asked for each individual by type of ownership. Information is gathered for assets held jointly with spouse and for assets held individually. There is some evidence that nonreporting of ownership is higher for assets held by one individual as compared to assets held jointly with others [Ferber and Frankel, 1978]. Asking respondents by type of ownership directly may tend to reduce differential nonresponse between assets held in own name and jointly with others. In addition, collecting income and asset values by type of ownership rather than one total amount for an asset may tend to give more accurate income and asset value amounts. Because both asset income and asset value are collected, information about asset income can be used for assessing the reasonableness of the asset value data and for imputing missing values.

Finally, "callback items" have been introduced for critical questions concerning asset values. Callback items are designed to reduce nonreporting of income and asset amounts. For selected items, when a respondent answers "don't know," the interviewer reads a statement on the importance of the information requested and the respondent is asked whether it would be possible to call back later for an estimate of the amount. For respondents agreeing to be called back, a "reminder card" is provided to the respondent with the requested information checked. The interviewer telephones the respondent at an agreed upon time to obtain the missing information which is entered in a section of the questionnaire reserved for callback amounts. The impact of the callback system is to reduce the incidence of missing information. Since callbacks involve added respondent and interviewer burden, only a limited number of items can be identified as callback items. To supplement the callback procedure, special instructions have been included for other important questions which instruct interviewers to probe for an estimate before accepting a "don't know" response. In this way, interviewers are alerted to key items for which they should give special effort to obtain estimates.

An important issue for surveys which measure wealth is population coverage. Evidence suggests that wealth holdings are concentrated. Studies have estimated that the top 5 percent of household hold approximately 30 to 50 percent of net wealth [Wolff, 1983; Greenwood, 1983; and Smith, 1983]. In addition, holdings of certain assets such as stocks are even more highly concentrated with the top 5 percent holding approximately 70 percent of this asset. As a result, the normal SIPP area frame sample has a limited coverage of the top wealthholders. This problem was noted by the Census Advisory Committee on Population Statistics which recommended that "the Bureau explore the opportunity to augment the SIPP sample design every 5 years to oversample the upper end of the income and wealth distribution, where special effort would be directed to producing a reliable indication of the entire distribution of wealth" [U.S. Bureau of the Census, 1983].

The Census Bureau has done some preliminary work to explore the methods which might be used to obtain data files in which top wealthholders are adequately represented. One possibility is to use IRS data files to develop a list frame sample of top income holders. An alternative approach is to use estate tax data to estimate wealth distribution of top wealthholders. Using estate tax data in conjunction with SIPP data has the potential to improve population coverage. At this time, however, the work in this area is very preliminary. In any case, the probable limitations of SIPP wealth estimates and the need to improve population coverage have been recognized by the Census Bureau and the work to improve the estimates continues.

III. Data and Results

A. Nonresponse Rates for Asset Ownership

Ownership of assets is established in the first wave of SIPP and an asset roster is constructed for each respondent. This roster is verified and updated in subsequent waves. The amounts of assets held are not systematically covered until the fourth wave topical module. In Table 1, the level of missing information on asset ownership are shown. Respondents can answer that they "don't know" if they own a specific asset type or they can refuse to answer. The nonresponse rate for all asset types is low at 1.4 percent of all persons asked about asset ownership. For specific asset types, the rates differ and range from 0.9 percent for rental property and royalties to 2.2 percent for certificates of deposit.[4] When the nonresponse rates are decomposed, the refusal rates are generally higher than the "don't know" rates. This result is true of self respondents. In general, self respondents have higher refusal than "don't know" rates. However, as would be expected, the frequency of "don't know" responses for asset ownership is greater for proxy respondents which range from 0.3 to 1.7 percent, as compared to self respondents which range from 0.1 to 0.3 percent. The refusal rates for both types of respondents were similar (ranging from 0.7 to 1.5 percent). As a result of higher "don't know" rates, proxy respondents had somewhat higher overall nonresponse rates. For both types of respondents, however, the absolute level of the nonresponse rates are low.

To examine the asset ownership information further, nonresponse rates by demographic and socioeconomic characteristics are shown in Table 2. Several patterns emerge. Nonresponse rates for each asset type are approximately the same between sex or race groups. Some of the rates, however, are significantly different by age and education levels. The nonresponse rates for each asset type increase with the age of the respondent. For respondents in the less than 25 and 25 to 34 age groups, nonresponse rates range between 0.3 to 1.2 percent; for the 35 to 44 age group, rates range from 0.9 to 2.0; for the 45 to 54 and 54 to 64 age groups, rates range from 1.1 to 2.8; and for the 65 and older age group, rates range from 1.3 to 3.5 percent. While the rates for the oldest age group are significantly higher than for the youngest group for each asset type, it should be noted that even the highest rates (3.4 and 3.5 percent for money market deposit accounts and certificates of deposit, respectively) are relatively low.

Nonresponse rates also differ by education level. The nonresponse rates for ownership of several asset types is higher for college graduates than for respondents who did not complete high school or who completed some college. Nonresponse rates for college graduates range from 1.3 to 2.6 percent, while the rates for respondents with less than high school or some college education levels range from 0.5 to 1.8 percent. Differences in rates between these groups are significant for each detailed asset types, except saving accounts and interest earning checking accounts.

B. Asset Ownership Patterns

Analyzing the SIPP asset ownership results, several patterns emerge. The frequency of asset ownership is shown in Table 3. The most frequently held assets were savings accounts and home ownership with 65.7 and 64.1 percent of households reporting ownership, respectively. The ownership of the remaining assets was reported by approximately 20 percent of households, with rental properties and royalties reported by 10 percent or less of households. Of interest are the asset types which became newly available after 1982 as a result of deregulation in the banking industry. Approximately 23 percent of the households reported ownership of one or more interest earning checking accounts, while 17 percent of the households reported ownership of money market deposit accounts.

The SIPP wave 1 results for selected assets are compared to asset ownership data from other surveys in Table 4. In general, the ownership estimates obtained in wave 1 of SIPP are similar to estimates derived from other sources. The only major difference occurs in savings accounts. The Consumer Credit Survey and the ISDP results show approximately 75 percent of households reporting ownership of savings accounts, while SIPP results show approximately 66 percent of the households owning such accounts. A plausible explanation is that savings accounts are highly substitutable with the newly available assets. Individuals have an incentive to switch from savings accounts to new accounts for the greater liquidity available with interest earning checking accounts and for the higher interest rates available with money market deposit accounts. The result is a negative impact on the percent of households owning savings accounts.

C. Asset Amount Response

An important feature of the SIPP design is to have income recipiency and asset ownership questions precede income and assest amounts questions. Once income recipiency is established for a respondent, questions on amount of income are asked by income type. Income from assets, e.g. interest, dividends and rental income, are covered in each wave. This section focuses on the reporting of property income and asset amounts in the first wave of SIPP.

The questions on asset income are divided by type of ownership, that is, assets held jointly with spouse, and assets in own name or held jointly with others. Persons identified as holding an asset are asked questions concerning the amount of income received. If respondents cannot provide the amount of interest earned, the average balances of accounts are asked in order to impute the interest income received. A callback is provided for the balance amount if the person does not know the amount at the time of the interview but can provide an estimate later. Unlike the ownership items, questions on interest amounts refer to grouped assets. In particular, assets held with financial institutions (savings accounts, money market deposit accounts, certificates of deposit, and interest earning checking accounts) are grouped together and the total interest earned is asked. Similarily, other interest earning assets (money market funds, U.S. Government securities, municipal or corporate bonds, and other interest earning assets) are grouped and interest income from those sources is covered.

The reporting of interest income is summarized in Table 5. On the average, two-thirds of respondents reported an amount of interest earned. Approximately 15 to 20 percent of respondents did not know the interest income amount, but provided an estimate of the total balance in the accounts. In general, over 80 percent of respondents gave the amount of interest or the balance in the accounts. Less than 11 percent did not know the amount of interest and did not provide a balance amount. Only 5 to 8 percent of respondents refused to report the interest income.

These patterns of reporting did not differ by type of ownership, that is, joint with spouse versus in own name or with others. However, the frequency of persons who refused or did not provide an interest/balance amount was lower on average for assets at financial institutions (11.7 percent) than other interest earning assets (17.1 percent).

Of interest are the cases which did not know the amount of interest earned and which were asked the average balance held for imputation purposes. Over 5,900 respondents were asked the balance of accounts held with financial institutions, and 455 were asked the balance of other interest earning accounts. For assets held with financial institutions, approximately 78 percent of the respondents which reported they did not know the interest income earned were resolved using the average balance reported; for other interest earning assets, 67 percent of the "don't know" cases reported a balance amount. Only 6 percent of the respondents refused to report an amount.

IV. Conclusion

The SIPP will provide annual estimates of wealth. Information on asset and liabilities is useful for many types of analyses, including program eligibility simulation studies and measurements of the distribution of wealth. In this paper, significant features of the survey were presented along with some preliminary results from SIPP. Several design features of SIPP, including longitudinal nature of the survey design, separating asset ownership and asset value questions, updating asset roster each interview, callbacks, probe instructions, and a missing wave section, are likely to have a positive effect on the reporting of asset ownership. In addition, asset and liability coverage is comprehensive. Results from the first wave of SIPP show that nonresponse rates for asset ownership are low. In addition, frequency of ownership patterns are reasonable and the results are comparable to findings from other surveys. Preliminary results also show some indication of improvements in nonresponse rates for items on asset amounts.

FOOTNOTES

1 When comparing CPS income aggregates to independent benchmark estimates, CPS captures 97.4 percent of independent wage and salary totals. However, the CPS only covers 41.5 and 44.1 percent of interest and dividend totals, 69.4 percent of SSI totals, and 42.3 percent of worker's compensation totals.

2 Persons who move individually or in groups are followed if they relocate within 100 miles of any SIPP PSU. Persons who move into institutions are not interviewed if they are institutionalized.

3 Home equity, automobiles, and life insurance information were collected in wave 2, primarily to evaluate government program eligibility.

4 Significance tests were performed on the differences between rates using the formula

$$\frac{(P_1-P_2)}{\frac{P_1(1-P_1)}{N_1} * F_1 + \frac{P_2(1-P_2)}{N_2} * F_2}$$

where P_1 and P_2 are the proportion of nonresponses, F_1 and F_2 are sample design factors, and N_1 and N_2 are the number of sample cases for each proportion. Differences noted in the text are significant at the 95 percent confidence level.

REFERENCES

[1] David, Martin H., "Technical, Conceptual, and Administrative Lessons of the Income Survey Development Program (ISDP)-Papers Presented at a Conference," Social Science Research Council, Washington, D.C., 1982.

[2] Ferber, Robert, John Forsythe, Harold Guthrie, and E. Scott Maynes, "Validation of Consumer Financial Characteristics: Common Stock," Journal of the American Statistical Association, Volume 64, Number 326, June 1969, pp. 415-432.

[3] Ferber, Robert and Matilda Frankel, "Evaluation of the Reliability of the Net Worth Data in the 1979 Panel: Asset Ownership on Wave 1," Survey Research Laboratory, University of Illinois, August 1981.

[4] Ferber, Robert and Matilda Frankel, "The Collection, Measurement, and Evaluation of Savings Account Reports," Survey Research Laboratory, University of Illinois, March 1978.

[5] Greenwood, Daphne, "An Estimation of U.S. Family Wealth and Its Distribution From Microdata, 1973," The Review of Income and Wealth, Series 29, Number 1, March 1983.

[6] Nelson, Dawn, David McMillen, and Daniel Kasprzyk, "Overview of the Survey of Income and Program Participation," U.S. Bureau of the Census, December 1983.

[7] Pearl, Robert and Matilda Frankel, "Composition of the Personal Wealth of American Households at the Start of the Eighties," Survey Research Laboratory, University of Illinois, presented at the American Economic Association meetings, December 1981.

[8] Pearl, Robert, Matilda Frankel, and Richard Williams, "The Effects of Missing Information on the Reliability of Net Worth Data From the 1979 ISDP Research Panel," Survey Research Laboratory, University of Illinois, May 1982.

[9] Projector, Dorothy S. and Gertrude S. Weiss, "Survey of Financial Characteristics of Consumers," Board of Governors of the Federal Reserve System, August 1966.

[10] Radner, Daniel B. and Denton R. Vaughan, "The Joint Distribution of Wealth and Income For Age Groups, 1979," Social Security Administration, Office of Research and Statistics, Working Paper 33, March 1984.

[11] Schwartz, Marvin, "Trends in Personal Wealth, 1976-1981," Statistics of Income Bulletin, Volume 3, Number 1, Summer 1983.

[12] Smith, James D., "Recent Trends in the Distribution of Wealth: Data, Research Problems and Prospects," presented at C.V. Starr Center Conference on International Comparisons of the Distribution of Household Wealth, New York University, November 1983.

[13] Smith, James D. and Stanton Calvert, "Estimating the Wealth of Top Wealth-Holders From Estate Tax Returns," 1965 Proceedings of the Business and Economic Statistics Section, American Statistical Association, pp. 248-265.

[14] U.S. Bureau of the Census, "Minutes and Report of Committee Recommendations, Census Advisory Committee on Population Statistics," April 8, 1983.

[15] U.S. Bureau of the Census, "Money Income of Families and Persons in the United States: 1980," Current Population Reports, Series P-60, No. 132, U.S. Government Printing Office, Washington, D.C., 1982.

[16] Vaughan, Denton R., T. Cameron Whiteman, and Charles A. Lininger, "The Quality of Income and Program Data in the 1979 ISDP Research Panel: Some Preliminary Findings," presented at the Social Science Workshop on the Survey of Income and Program Participation, Baltimore, Maryland, December 1983.

[17] Wolff, Edward N., "The Size Distribution of Household Disposable Wealth in the United States," The Review of Income and Wealth, Series 29, Number 2, June 1983, pp. 125-146.

[18] Ycas, Martynas A. and Charles A. Lininger, "The Income Survey Development Program: Design Features and Initial Findings - Paper 1," presented at Social Science Research Council Conference, Washington, D.C., October 1982.

Table 1.--Percent Missing Information For Asset Ownership By Type of Respondent, SIPP

Asset type	Total			Proxy			Self		
	Total	Don't know	Refusal	Total	Don't know	Refusal	Total	Don't know	Refusal
Total	1.4	0.4	1.0	1.9	1.0	0.9	1.2	0.2	1.0
Savings Accounts ...	1.7	0.4	1.3	2.1	0.9	1.2	1.4	0.1	1.3
Money Market Deposit Accounts..	2.1	0.7	1.4	2.8	1.6	1.2	1.7	0.3	1.4
Certificates of Deposit	2.2	0.8	1.4	3.0	1.7	1.3	1.8	0.3	1.5
Interest Earning Checking Accounts	1.9	0.6	1.3	2.8	1.5	1.3	1.6	0.3	1.3
Other Interest Earning Assets[1] ..	1.4	0.4	1.0	1.7	0.9	0.8	1.2	0.2	1.0
Stocks and Mutual Funds	1.2	0.4	0.8	1.6	0.8	0.8	1.1	0.2	0.9
Rental Property	0.9	0.1	0.8	1.0	0.3	0.7	0.9	0.1	0.8
Royalties	0.9	0.2	0.7	1.1	0.4	0.7	0.9	0.1	0.8
Other Investments ..	1.3	0.5	0.8	1.6	0.9	0.7	1.1	0.3	0.8

[1]Consist of money market funds, U.S. Government securities, municipal and corporate bonds, mortgages, and savings bonds.

TABLE 2.--ASSET OWNERSHIP NONRESPONSE RATES BY CHARACTERISTIC, WAVE 1 SIPP

Characteristics	Asset type: Total	Savings accounts	Money market deposit accounts	Certificates of deposit	Interest earning checking accounts	Other interest earning assets	Stocks or mutual funds	Rental property
Total	1.4	1.6	2.1	2.2	2.0	1.4	1.2	0.9
SEX								
Male	1.5	1.7	2.2	2.3	2.1	1.5	1.3	1.0
Female	1.3	1.6	2.0	2.0	1.8	1.3	1.2	0.8
RACE								
White	1.5	1.6	2.1	2.2	1.9	1.4	1.3	0.9
Black	1.2	1.9	2.3	2.5	2.6	1.1	0.9	0.6
Other	0.7	0.9	1.4	1.4	0.9	0.7	0.4	0.3
AGE								
Less than 25	0.6	0.7	0.9	0.9	1.0	0.5	0.4	0.3
25 to 34	0.7	0.8	1.0	1.1	1.2	0.7	0.7	0.4
35 to 44	1.4	1.5	2.0	2.0	1.9	1.4	1.4	1.0
45 to 54	1.9	2.4	2.7	2.9	2.5	1.8	1.6	1.1
55 to 64	2.0	2.3	2.8	2.8	2.5	2.1	1.8	1.3
65 and over	2.5	2.7	3.4	3.5	3.0	2.6	2.1	1.7
EDUCATION LEVEL								
Less than high school	1.0	1.5	1.8	1.7	1.6	0.9	0.7	0.6
High School	1.6	1.8	2.2	2.3	2.1	1.5	1.3	1.0
Some college	0.9	1.0	1.3	1.4	1.2	1.0	0.8	0.5
College graduate..	2.0	1.8	2.4	2.6	2.2	2.1	2.1	1.4

Table 3.--Percent of House holds Reporting Ownership Asset Types[1]

Asset type	SIPP Wave 1
Savings accounts	65.7
Money market deposit accounts	17.1
Certificates of deposit	19.0
Interest earning checking accounts ...	23.1
Other interest earning assets[2]	18.1
Stocks or mutual fund shares	20.9
Rental property	10.2
Royalties	1.4
Home ownership	64.1

1 Preliminary SIPP wave 1 results.

2 Includes money market funds, U.S. Government securities, municipal or corporate bonds, mortgages, and U.S. Savings Bonds.

TABLE 4.--PERCENT OF HOUSEHOLDS REPORTING OWNERSHIP OF SELECTED ASSETS

ASSET TYPE	SIPP WAVE 1 (1983)	ISDP (1979)	CONSUMER CREDIT SURVEY (1977)	SURVEY OF FINANCIAL CHARACTERISTICS OF CONSUMERS (1962-63)
SAVINGS ACCOUNTS	65.7	73.8	76.9	59.0
CERTIFICATES OF DEPOSIT	19.0	15.7	13.6	-
STOCKS OR MUTUAL FUND SHARES .	20.9	19.6	24.7	16.0
RENTAL PROPERTY	10.2	13.6	-	-
HOME OWNERSHIP	64.1	60.5	61.0	57.0
SAMPLE SIZE	19,878	6,992	2,563	2,557

TABLE 5.--DISTRIBUTION OF INTEREST INCOME REPORTING

ASSET TYPE	TOTAL	INTEREST INCOME REPORTED	DON'T KNOW: ASSET AMOUNT REPORTED	DON'T KNOW: NO ASSET AMOUNT REPORTED	REFUSAL
TOTAL					
ASSETS AT FINANCIAL INSTITUTIONS ..	100.0	66.8	21.5	7.2	4.5
OTHER INTEREST EARNING ASSETS	100.0	67.6	15.3	10.2	6.9
JOINT WITH SPOUSE					
ASSETS AT FINANCIAL INSTITUTIONS ..	100.0	68.0	20.0	6.5	5.5
OTHER INTEREST EARNING ASSETS	100.0	60.4	13.6	10.8	8.1
OWNERSHIP IN OWN NAME					
ASSETS AT FINANCIAL INSTITUTIONS ..	100.0	66.0	22.5	7.6	3.9
OTHER INTEREST EARNING ASSETS	100.0	69.2	15.8	9.2	5.8

THE WEALTH AND INCOME OF AGED HOUSEHOLDS

Daniel B. Radner, Social Security Administration

I. Introduction*

The economic status of the aged population has been a topic of interest to researchers for some time, and there has been particular interest in recent years. In this paper we are interested in the economic resources of aged households. This topic is examined using data on both income and wealth. Estimates are presented and compared for all aged households and for aged households with both low income and low wealth. Socioeconomic characteristics, mean and median amounts, and the composition of wealth and income are discussed.

Many different indicators of economic well-being have been used by researchers. Income, specifically money income before taxes, is the most frequently used measure of economic well-being. Money income before taxes is also the definition of income used here. Of course, this definition is far from an ideal measure of economic well-being. An important exclusion is noncash income, which is a significant source of economic resources for most groups of the population. Also, by using pre-tax income, the resources available to the unit can be distorted. These deficiencies can affect comparisons between nonaged and aged households.

Measures of economic well-being that are confined to income omit the wealth of the unit, although income from assets ordinarily is included in income. A relatively new data base, the 1979 Income Survey Development Program (ISDP) file, makes the examination of both income and wealth possible. 1/ Wealth and income data can be used together in several different ways. In this paper a simple approach is used--income and wealth are used as a two-dimensional classification. This approach does not spread the wealth out over the expected lifetime of the unit as some methods (e.g., Weisbrod and Hansen 1968) do, but is concerned with a much shorter time horizon. In this application, the liquidity of the assets held can be very important. Although the most comprehensive definition available (net worth) is used here as the measure of wealth, the composition of wealth is examined so that liquidity can be taken into account.

In a recent paper, Radner and Vaughan (1984) examined the joint distribution of wealth and income for various age groups and, using several different definitions, looked at the proportion of households with both relatively low wealth and relatively low income. They showed that, despite the fact that average net worth is fairly high for aged households, the joint distribution of income and wealth is such that the percentage of aged households (age 65 or older) with both relatively low income and relatively low net worth exceeds the percentage for each age group in the 25-64 age range.

Aged households identified using one of those definitions of low income and low wealth are examined further in this paper. The particular definition used is that the household had to be in the bottom 20 percent of all households ranked by size of income and in the bottom 40 percent of all households ranked by size of net worth. For both income and net worth the rankings were based on amounts that had been adjusted for size of household using an equivalence scale derived from the U.S. poverty lines.

All estimates shown in this paper are on a household basis. For convenience it is assumed that the income and wealth of the household are resources for all members of the household and only for members of the household. Thus, relatives of persons in the household who do not live in the household (e.g., parents or children) are assumed to have no claim on the resources of the household, and the household is assumed to have no claim on other resources. Some of the estimates shown in this paper take into account the number of persons in the household.

Because the characteristics and economic situations of aged households can differ substantially by age, estimates are shown for detailed age groups within the aged group. Space limitations and small sample sizes restrict the age detail that can be shown.

Section II briefly describes the data used and presents definitions of important concepts used in the paper. Section III provides a brief overview of the income and wealth of households in different age groups. In Section IV characteristics of all aged households and of aged households with both relatively low income and relatively low wealth are presented and discussed. A summary and conclusions are presented in Section V.

II. Data and Definitions

The data used in this paper are from the 1979 ISDP file (Radner and Vaughan 1984). The sample was nationally representative and both the low and high ends of the income distribution were oversampled slightly. Detailed information on income, assets, debts and socioeconomic characteristics was obtained. The data used in this paper are primarily from the fifth wave of this multi-wave panel. The fifth wave, which has most of the data on wealth, contains observations on about 6,900 households resulting from interviews in January, February, and March 1980.

The ISDP income data suffer from the underreporting which is common to household surveys. However, overall the ISDP data appear to be better than the income data in the Current Population Survey (CPS). The quality of the wealth data is difficult to assess, although it is known that misreporting and nonresponse are substantial problems. While there is some evidence of marked increases in estimates of asset ownership compared to other surveys, unfortunately item nonresponse on asset values (for assets other than owner-occupied housing and vehicles) is quite high. Thus, very substantial proportions of the final asset value aggregates stem from values assigned using "hot deck" imputation. This is a problem that should be kept in mind in interpreting the estimates shown. The post-imputation estimates of net worth and of most asset types used in this paper suffer from substantial underreporting. Financial assets appear to show the highest percentage of underreporting. However, a rough comparison of survey aggregates to independent control aggregates suggests results that are similar to the 1962 Survey of Financial Characteristics of Consumers (Projector and Weiss 1966). It should be noted that the ISDP estimates of the extreme upper tail of the net worth distribution show a far lower share of net worth than is shown by other sources of wealth data. This "absence" of the extreme upper tail results at least in part from an emphasis in the survey on obtaining data for low- and middle-income units.

In general the demographic concepts used in this paper are the same as or similar to those employed by the Bureau of the Census in its Annual Demographic Supplement to the CPS. 2/ The data presented pertain to the civilian noninstitutional population living in the 50 states and the District of Columbia at the dates of interview. All estimates are on a household basis. Age classifications are based on the age of the household reference person (householder) at his or her most recent birthday. The householder is the person (or one of the persons) in whose name the home is owned or rented.

Income is defined on a before-tax basis and is presented at annualized rates; that is, as the measured three-month value times four. All money income received by household members during the three months preceding the month of interview is covered, including one-time or lump sum payments such as life insurance proceeds and gifts. Total money income consists of the sum of earnings, property income, OASDI, SSI, pensions, and other income. 3/ Earnings includes wages and salaries, net income from farm self-employment, and "draw" from nonfarm self-employment. 4/ Property income includes interest, dividends, rent, royalties, and income from estates and trusts. OASDI consists of social security and railroad retirement income. SSI consists of Supplemental Security Income, including both state and federal amounts. Pensions are the sum of government and private pensions. Government pensions consists of U.S. civil service and military retirement pensions and state and local government pensions. Private pensions includes employer and union pensions and annuities. Other income includes unemployment compensation, worker's compensation, veterans' compensation and pensions, public assistance, educational benefits, alimony, assistance from relatives and friends, and lump sum payments.

The wealth concept used is net worth. Net worth consists of all assets less all debts covered by the survey. Net worth is defined to be wealth minus unsecured debt. All net worth components except home and vehicle equity are valued as of the end of 1979; those two components are valued as of mid-1979. 5/ Wealth is defined as the value of all assets covered by the survey less any debts secured by those assets. Several items sometimes included in estimates of wealth are not covered; for example, social security wealth, pension wealth, and trusts are excluded.

Wealth is the sum of the following items: home equity (owner-occupied); durable goods (equity in vehicles plus market value of household durables); business equity (nonfarm and farm); liquid financial assets (cash, checking accounts, passbook savings accounts, U.S. savings bonds); nonliquid financial assets (bonds, CD's, stocks, mutual fund shares); other assets (e.g., equity in other property, equity in nonactive business interest). Unsecured debt includes installment and noninstallment debt, unpaid medical bills, and educational loans.

III. General Patterns of Income and Wealth by Age of Householder

In this section an overview of the relationship between age and income and wealth is presented. Mean and median amounts are shown, and the dispersion of the income and wealth of aged households is discussed. Mean and median total money income and net worth are shown for age (of householder) groups in Table 1. 6/ 7/ Looking at income, these estimates show the familiar pattern of relatively low mean and median at young and old ages, with a peak in the 45-54 age group. The mean (median) for the 65 and over group is below the mean (median) for the youngest age group. Of course, the median is below the mean for every age group. The pattern for net worth is somewhat different. Mean amounts are low for the younger age groups, peak in the 55-64 group, and then decline. However, the mean (median) for the 65 and over group is far above the mean (median) for the three youngest age groups. In general, the median is about one half the mean in each age group. Within the detailed aged groups, both income and net worth decline as age increases, using both means and medians, with the 75 and over group substantially below the 65-69 group.

One way of looking at the mean amounts is to compute relative means. The relative mean is the mean for the group divided by the mean for all households. Thus, the relative mean income for households age 65 and over is \$13,070/\$21,260 = 0.61. Although households age 65 or over have a relative mean income that is only 61 percent of the overall mean, their relative mean net worth is 27 percent above the overall mean.

It is well-known that households with aged heads are, on average, smaller than households with nonaged heads. These differences in household size are adjusted for in this paper using an equivalence scale based on the U.S. poverty lines (U.S. Bureau of the Census 1981b, Table A-3). 8/ A one-person household is taken as the base, and the income and net worth of each household are divided by the appropriate scale value to obtain amounts adjusted for size of household. 9/ The relative means of the 35-54 age groups, which have relatively large households, show declines from the unadjusted estimates. The 55-64 and 65 and over groups, which have relatively small households, show increases. The relative mean income for the 65 and over group rises to 0.79, while the relative mean net worth increases to 1.59. In general, the relative means of aged households rise about 25 percent.

There is substantial dispersion that cannot be seen by just looking at the averages. In the distribution by size of income (after adjustment for size of household), substantial dispersion among quintiles is present for the 65 and over age group. For example, even though 32 percent of households are in the bottom income quintile, 11 percent appear in the top quintile. The distribution among net worth quintiles also shows substantial dispersion, with 24 percent in the bottom two quintiles and 34 percent in the top quintile.

IV. Characteristics of Aged Households

In this section all aged households and aged households with both relatively low income and relatively low net worth are discussed and compared. A household is defined to have relatively low income if it is in the bottom quintile (20%) in the overall (all ages) distribution of households by size of income. A household is defined to have relatively low net worth if it is in the bottom two quintiles (40%) in the overall (all ages) distribution of households by size of net worth. In both cases the amounts were adjusted for size of household. 10/ Households that have both low income and low net worth using these definitions will be called "LILNW" (Low Income and Low Net Worth) households.

Because the estimates have been adjusted using a scale based on the poverty lines, the dollar amount cutoffs that define the LILNW group differ by household size and, for one- and two-person households, by age of householder. For one-person households with householder age 65 or over, the upper (annualized) income bound for the LILNW group is $5,047, while the upper net worth bound is $9,875. For two-person households with householder age 65 or over, the income bound is $6,369, while the net worth bound is $12,462. The bounds are about nine percent higher if the householder is under age 65. 11/

Thirteen percent of all households are in the LILNW group. The percentages range from 9 percent for households age 55-64 to 21 percent for households under 25. Fifteen percent of all households 65 and over are in the LILNW group; the percentage is 11 in the 65-69 group, 15 in the 70-74 group, and 19 in the 75 and over group. 12/ Thus, substantial percentages of the aged groups have both relatively low income and relatively low net worth. We have seen that 24 percent of households age 65 and over are in the bottom two net worth quintiles, while 32 percent are in the bottom income quintile. Since only 15 percent are in both groups, it can be seen that about two thirds of those with low net worth also had low income (15/24), and about one half of those with low income also had low net worth (15/32).

We are interested here in who the aged LILNW households are and what resources they have. Table 2 shows the composition of the LILNW group compared to all households for the 65 and over and 75 and over age groups. Looking at sex of householder, 46 percent of all households 65 and over have a female householder, but 75 percent of the LILNW households have a female householder. For the 75 and over group, 86 percent of the LILNW households have a female householder. For household size, the LILNW group shows substantially more one-person households than the population as a whole in the age group (76 percent compared to 44 percent in the 65 and over group). The LILNW group shows substantially fewer married spouse present householders (11 percent compared to 43 percent in the 65 and over group) and more widowed, divorced, and other householders. Looking at these three characteristics together, households of size one with a widowed female householder comprise 49 percent of the LILNW 65 and over group, but only 27 percent of the total in that age group. The corresponding percentages for the 75 and over group are 68 percent in the LILNW group and only 37 percent overall.

Table 3 shows the composition of total income, mean amounts of income types, and the percentage of households with various types of income for all households in the 65 and over and 75 and over groups and for LILNW households in those age groups. Looking at all households 65 and over, OASDI accounts for about one third of total income, while earnings and property income each constitute about one fourth, and pension income accounts for about one tenth. The composition of income differs markedly between LILNW and all households. Looking at the 65 and over group, OASDI benefits are far more important in the LILNW group (about two thirds of total income), although the mean amount of benefits in that group is far below the mean for all households. SSI is also more important in the LILNW group (12 percent compared to 1 percent). All other income sources are less important in the LILNW group, with earnings (9 percent compared to 25 percent) and property income (2 percent compared to 23 percent) showing large differences. The patterns are similar for the 75 and over group. Mean total income for the LILNW group is only 29 percent of the mean for all households in the 65 and over group and 32 percent in the 75 and over group. Because median income is quite close to mean income in the LILNW group, but not for all households, the ratio of the LILNW median to the overall median is higher--42 percent for the 65 and over group and 48 percent for the 75 and over group. For both mean and median, there is less difference between the 65 and over and 75 and over groups for the LILNW group than for all households. Panel C of Table 3 shows the percentage of the group that received each type of income. Households in the LILNW group receive SSI much more often (32 percent compared to 11 percent for the 65 and over group), and earnings, property income, and pensions much less often than all households in the age group. The percentages receiving OASDI (93-94 percent for the 65 and over group) show little difference.

Table 4 shows the composition of net worth for LILNW households and all households in the 65 and over and 75 and over age groups. For all households 65 and over, home equity and financial assets each account for about one third of net worth, while durable goods constitute about one tenth. In contrast, in the LILNW group 65 and over, home equity is about one fifth of net worth and financial assets are about one third, but durable goods constitute almost three fifths. The LILNW mean of each asset is far below the mean for all households in the age group. The LILNW group shows larger shares for liquid financial assets and durable goods, negligible shares for nonliquid financial assets, business equity, and other assets, and a lower share for home equity. Mean and median amounts of net worth are very low in the LILNW group, even below the income amounts. If durable goods are excluded from the mean (on the grounds that that type of asset cannot easily be liquidated or borrowed against), then the mean amount is $1,260 (including home equity).

Panel C of Table 4 shows the percentage of households with each type of asset or debt. Home equity, nonliquid financial assets, business equity, and other assets are held by few households in the LILNW group. Durable goods and liquid financial assets are held by high percentages. Fewer households in the LILNW group have unsecured debt. Differences between the 65 and over and 75 and over groups are relatively small.

V. Summary and Conclusions

We have seen that, on the average, aged households have relatively low incomes and relatively high net worth. However, because there is substantial dispersion in these amounts, the averages do not tell the whole story. For example, at the end of 1979 about 2 1/2 million aged households had both relatively low income (mostly OASDI and SSI) and relatively low net worth (mostly durable goods). However, there also are aged households with substantial amounts of both income and net worth. Thus, general statements about how well off "the aged" are can frequently be misleading and should be interpreted with caution.

FOOTNOTES

*The author is greatly indebted to Sharon Johnson, who prepared the estimates, and to Benjamin Bridges and Denton Vaughan for their helpful comments.

1/ In this paper we were not able to incorporate the data on taxes and noncash income that were collected in the ISDP.

2/ See U.S. Bureau of the Census (1981a) for the CPS definitions.

3/ In the case of income-producing assets we include both the income (in income) and the asset value (in net worth).

4/ Net income fom nonfarm self-employment was not obtained in Wave 5 of the ISDP. The "draw" is the amount of salary or money taken out of the business for living expenses.

5/ Those two items were collected in Wave 2 of the ISDP for most households.

6/ The estimates in this paper differ slightly from those in Radner and Vaughan (1984) because these estimates reflect our approximate correction of errors that we believe exist in the Wave 5 ISDP public use file. Two errors relating to the processing of data from the 3-month-6-month property income recall test appear to have been made in the construction of household summary amounts of property income from amounts of specific income types. First, all of Rotation 3 was treated as a 6-month reference period group, whereas only one half of Rotation 3 had actually used the 6-month reference period. Second, for all of Rotation 3, royalty and estate and trust incomes were treated as though the 6-month reference period had been used; actually, a 3-month reference period was used for those income types for all households. These two errors resulted in an underestimate of about $3 1/2 billion in property income and total income on the household records. We were able to make only an approximate correction of these errors in the estimates shown in this paper.

7/ The estimates in this paper exclude a few observations with negative household income.

8/ The choice of this scale was arbitrary. However, this is a familiar scale that is much less extreme than a per capita adjustment. The extensive literature on equivalence scales contains many estimated scales that could have been used. In some cases, the estimates shown in this paper could be sensitive to the choice of a scale.

9/ The scale values used are: 1 person (under 65), 1.024; 1 person (65+), 0.943; 2 persons (under 65), 1.322; 2 persons (65+), 1.190; 3 persons, 1.568; 4 persons, 2.009; 5 persons, 2.379; 6 persons, 2.687; 7 persons or more, 3.329.

10/ It should be noted that these definitions are somewhat arbitrary. The bottom two quintiles were used instead of the bottom quintile for net worth because the implied net worth amounts in the bottom quintile were extremely small. Looking at these low income and low net worth households does not require a completely specified ranking of households according to economic well-being. Also, we do not claim that all low income and low net worth households are worse off than all other households.

11/ The bounds for other household sizes can be derived using the scale values shown in footnote 9. The income bounds are about 45 percent above the 1979 poverty lines.

12/ Using estimates unadjusted for household size, 20 percent of aged households were in the LILNW group, compared to 13 percent for all ages (Radner and Vaughan 1984, Table 16.) For comparison, in 1979, about 12 percent of all households and 18 percent of aged households were poor (U.S. Bureau of the Census 1982, Table 22).

REFERENCES

Projector, Dorothy S. and Gertrude S. Weiss (1966), Survey of Financial Characteristics of Consumers, Board of Governors of the Federal Reserve System.

Radner, Daniel B. and Denton R. Vaughan (1984), "The Joint Distribution of Wealth and Income for Age Groups, 1979," ORS Working Paper No. 33, Office of Research, Statistics, and International Policy, Social Security

Administration (March).

U.S. Bureau of the Census (1981a), Money Income of Families and Persons in the United States: 1979, Current Population Reports, Series P-60, No. 129, Washington.

U.S. Bureau of the Census (1981b), Characteristics of the Population Below the Poverty Level: 1979, Current Population Reports, Series P-60, No. 130, Washington.

U.S. Bureau of the Census (1982, Characteristics of Households and Persons Receiving Selected Noncash Benefits: 1980 (With Comparable Data for 1979), Current Population Reports, Series P-60, No. 131, Washington.

Weisbrod, Burton A. and Lee W. Hansen (1968), "An Income-Net Worth Approach to Measuring Economic Welfare," American Economic Review, Vol. LVIII, No. 5 (December) pp. 1315-29.

Table 1.--Mean and Median Income and Net Worth, by Age of Householder, 1979

Age of Householder	Sample Cases	Weighted Number of Households (Thousands)	Annualized Income ($) Mean	Annualized Income ($) Median	Net Worth ($) Mean	Net Worth ($) Median
All Ages	6,922	82,211	21,260	16,500	62,430	25,770
Under 25	621	6,613	14,150	12,860	8,880	3,450
25-34	1,441	19,272	20,160	18,000	24,520	12,110
35-44	1,089	14,014	25,390	20,430	64,950	29,820
45-54	1,049	12,975	28,400	24,180	79,120	39,180
55-64	1,155	12,722	25,500	19,330	105,740	51,440
65 and over	1,567	16,614	13,070	8,630	79,380	38,720
65-69	489	5,570	15,560	9,750	90,040	47,300
70-74	411	4,818	12,790	9,040	85,480	45,440
75 and over	667	6,226	11,050	7,460	65,130	28,410

NOTE: Dollar amounts are rounded to the nearest $10.

Table 2.--Percentage Composition of Aged Households, 1979

Type of Household	Age of Householder: 65 and over, Total	65 and over, Low Income and Low Net Worth	75 and over, Total	75 and over, Low Income and Low Net Worth
Total	100	100	100	100
Sex of Householder				
Female	46	75	54	86
Male	54	25	46	14
Household Size				
1 Person	44	76	50	88
2 Persons or more	56	24	50	12
Marital Status of Householder				
Married, Spouse Present	43	11	36	7
Widowed	41	61	55	75
Divorced	6	13	2	4
Other	9	15	9	14
Household Size 1, Householder Widowed Female	27	49	37	68
Sample Cases	1,567	402	667	201
Weighted Number of Households (thousands)	16,614	2,562	6,226	1,213

Table 3.--Percentage Composition of Income, Mean Incomes, and Percent Receiving Each Income Type, 1979

	Age of Householder: 65 and over		75 and over	
	Total	Low Income and Low Net Worth	Total	Low Income and Low Net Worth
	A. PERCENTAGE COMPOSITION			
Total	100	100	100	100
Earnings	25	9	17	3
Property	23	2	26	2
OASDI	34	68	39	75
SSI	1	12	2	10
Pension	11	4	9	3
Government	6	4	5	2
Private	5	1	4	1
Other	6	5	7	6
	B. MEAN AMOUNTS ($)			
Total	13,070	3,750	11,050	3,490
Earnings	3,240	330	1,890	120
Property	3,030	60	2,900	60
OASDI	4,400	2,540	4,310	2,620
SSI	190	450	230	360
Pension	1,450	170	980	120
Government	770	130	540	80
Private	680	40	440	40
Other	760	200	740	220
Addendum:				
Median Total Income ($)	8,630	3,630	7,460	3,580
	C. PERCENT RECEIVING EACH INCOME TYPE			
Total	100	100	100	100
Earnings	28	15	21	11
Property	77	36	80	45
OASDI	94	93	97	95
SSI	11	32	13	28
Pension	39	12	33	14
Government	15	6	13	6
Private	28	6	22	8
Other	19	27	22	24

NOTE: Dollar amounts are rounded to the nearest $10.

Table 4.--Percentage Composition of Net Worth, Mean Amounts of Assets, and Percent Holding Each Type of Asset, 1979

	Age of Householder: 65 and over		75 and over	
	Total	Low Income and Low Net Worth	Total	Low Income and Low Net Worth
	A. PERCENTAGE COMPOSITION			
Net Worth	100	100	100	100
Wealth	101	108	101	111
Home Equity	32	18	30	9
Financial Assets	35	31	44	43
Liquid	10	31	12	42
Nonliquid	25	1	32	1
Business Equity	7	0	5	0
Other Assets	18	0	13	0
Durable Goods	9	58	8	59
Unsecured Debt	1	8	1	11
	B. MEAN AMOUNTS ($)			
Net Worth	79,380	3,000	65,130	3,110
Wealth	79,930	3,240	65,530	3,460
Home Equity	25,110	550	19,480	290
Financial Assets	28,020	940	28,610	1,320
Liquid	8,020	920	7,560	1,290
Nonliquid	20,000	20	21,040	30
Business Equity	5,660	0	3,400	0
Other Assets	14,090	0	8,640	0
Durable Goods	7,060	1,740	5,410	1,840
Unsecured Debt	540	240	400	350
Addendum:				
Median Net Worth ($)	38,720	2,170	28,410	2,300
	C. PERCENT HOLDING EACH ASSET			
Net Worth	100	97	100	99
Wealth	99	97	100	99
Home Equity	66	8	59	4
Financial Assets	95	82	96	86
Liquid	95	82	96	86
Nonliquid	36	2	31	3
Business Equity	4	0	3	0
Other Assets	18	0	16	0
Durable Goods	98	96	99	99
Unsecured Debt	39	26	28	21

NOTE: Dollar amounts are rounded to the nearest $10.

USING SUBJECTIVE ASSESSMENTS OF INCOME TO ESTIMATE FAMILY EQUIVALENCE SCALES: A REPORT ON WORK IN PROGRESS*†

Denton R. Vaughan
Social Security Administration

This paper will present a short discussion of the estimation of family equivalence scales based on individuals' assessments of their own current income level or their subjective judgments about the income required to provide specified levels of satisfaction or to sustain a given standard of living.[1] A scale estimated on the basis of individuals' subjective assessments of their current income is presented and the results are discussed in terms of scales obtained by "traditional" methods and other "subjective-based" scales.

Techniques for establishing equivalence.

Traditional approaches.--A number of techniques have been used to estimate family size equivalence scales. Until recently, most approaches have relied either on some manifestation of consumer behavior that is interpreted as a welfare indicator or on scientific or quasi-scientific dietary standards.

The general class of Engel curve analyses falls into the first category and the most frequently used welfare indicator has been the percentage of income or expenditures devoted to food as affected by family size. As food has become a less dominant element of family budgets in Western Europe and the United States, and therefore less representative of the general population, Engel curve analyses have been extended to broader sets of consumption goods such as necessities (e.g., Sydenstriker and King 1921; Prais and Houthakker 1955; Watts 1967; Seneca and Taussig 1971). Other Engel curve applications have made use of adult consumption goods such as alcohol and tobacco (Nicholson 1949, 1976) or have analyzed variation in the percentage of income devoted to savings (U.S. Department of Labor 1948).

More recently, economists have attempted to define equivalence on the basis of demand equations estimated for the full set of consumption goods. In this approach, equivalence estimates are derived from variations in demand that are associated with family size and composition differences (Kakwani 1977; van der Gaag and Smolensky 1982; Danziger, van der Gaag, Smolensky and Taussig 1984).

Equivalence scales based on dietary adequacy take as a point of departure a recognized nutritional standard, such as provided by the U.S. National Research Council (NRC), and translate the standard into a per capita food plan appropriate to a given level of living. Since nutritional standards such as the one developed by the NRC incorporate variations by age and sex, their translation into food plans necessarily captures family composition effects (Peterkin 1976). By taking into account the interaction between per capita income, per capita food plan costs and the nutritional quality of diet (by family size), such dietary equivalence scales are said to capture economies of scale as well as compositional effects (e.g. Kerr and Peterkin 1976, p. 72-73). The equivalence scale incorporated in the official Federal poverty line is based in large part on an earlier version of the U.S. Department of Agriculture scale described by Peterkin. However, *ad hoc* adjustments were made to the ratios for families of one and two to account for the supposed higher fixed costs of operating small households (Orshansky 1965, p. 9).

Direct subjective measures. Beginning in the early 1970's, the economist van Praag and his colleagues at Leyden University in the Netherlands (van Praag 1971; van Praag and Kapteyn 1973; Kapteyn and van Praag 1976; Goedhart, Halberstadt, Kapteyn and van Praag 1977; and van Herwaarden, Kapteyn and van Praag 1977) and sociologists in the United States (Rainwater 1973 and 1974; Vaughan and Lancaster 1979; Dubnoff, Vaughan and Lancaster 1981) initiated the use of individuals' perceived income needs to estimate family equivalence scales. By the early 1980's a few additional U.S. social scientists had begun to exploit the "subjective" approach using newly available U.S. data bases (Colasanto, Kapteyn and van der Gaag 1984; Danziger, van der Gaag, Taussig and Smolensky 1984) in conjunction with Dutch economists then resident in the United States.

Actually, from a theoretical perspective the use of the term "subjective" to differentiate this class of measures from more orthodox approaches based on consumer behavior, and most particularly consumption choices, is somewhat misleading. In economics, behavior (consumer choice) is generally considered to be revelatory of subjective states, i.e., preferences are "revealed" by behavior, and much of the theorizing underlying the study of consumer behavior and welfare economics relies on primitive notions of subjective states. In the context of constraints, it is these states which essentially motivate, if not determine, behavior. So in an important sense what is novel about using subjective assessments to make inferences about preferences, utility, or welfare is not that in being subjective they are somehow set apart from the concerns of economists. Rather it is that the orthodox economist is more familiar, and therefore more comfortable, with employing behavior as an indicator of the underlying construct of welfare than with using the direct verbal representations of those constructs or states (Muellbauer 1980, p. 153; Pollak and Wales 1979, p. 219). Weighing in on the other side of the issue, however, Sen (1982, pp. 71-72) has noted that the "idea that behavior is the one real source of information [about preferences] is extremely limiting for empirical work and is not easy to justify in terms of the methodological requirements of our discipline [viz. economics]."

In any case, in the present context I will use the phrase direct subjective measures to at once highlight the theoretical relationship between this class of measures and those employed within a "revealed preference" framework while correctly stressing that they are based on consumers' perceptions or subjective assessments rather than on consumer behavior.

Types of direct subjective measures.

There are three basic types of direct subjective measures that have been used to estimate family equivalence scales. They are:

1) The income-subjective welfare metric.--This approach was pioneered and systematically exploited by van Praag and his associates (van Praag 1968, 1971; van Praag and Kapteyn 1973), but was also briefly experimented with by Andrews and Withey (1976, pp. 228-229, 378). Respondents are asked a

series of questions which permit estimation of the function which relates income to the welfare or utility derived from income as perceived by the individual. The parameters of the estimated function may then be used in a variety of applications, such as the construction of family size equivalence scales.

2) The income-living level approach. This technique relies on respondents' estimates of the money income amounts required to attain a given level of living (getting along, living comfortably, etc.). Systematic variation in the subjective income requirements associated with the various levels of living is used to model differences in underlying needs and to construct equivalence scales.

Rainwater (1974) was the first to obtain estimates for a range of alternative levels simultaneously. More recently the Expert Committee on Family Budget Revisions, convened by the Bureau of Labor Statistics, recommended a program of research to explore the feasibility of estimating a set of living level standards at the national level (Watts 1980). Research funded independently by the U.S. National Science Foundation (see Dubnoff and Strate 1984) is pursuing these and related issues.

3) Income rating scales. This method relies on obtaining the respondent's rating of his or her current income using a set of ranked categories (such as from "delighted" to "terrible") believed to extend over most, if not all, of the continuum of satisfaction or welfare. The sensitivity of such responses to family size was first demonstrated in 1979 (Vaughan and Lancaster) and the first equivalence scales per se were estimated using current income rating measures in 1981 (Dubnoff, Vaughan and Lancaster). This class of measures may be thought of as a subset of the income/welfare metric approach (Vaughan and Lancaster 1980).

Estimation of equivalence scales using income rating measures: An example.

This section of the paper reports on the preliminary stages of a project undertaken to evaluate the usefulness of subjective assessments for comparing differences in the level of economic welfare experienced by significant subgroups of the population. Developing equivalence estimates constitutes a major segment of this project. A number of the questions which will be addressed in the course of this work cannot be dealt with using the data set currently available, so the purpose of the current exercise is to give the reader the flavor of how one might address such issues on the basis of income rating scales.

The scale estimation procedure. The general approach is similar to that first employed by Dubnoff, Lancaster and the current author in an earlier paper (Dubnoff, Vaughan and Lancaster 1981). Essentially, a measure of satisfaction with current income is regressed on income, family size, and a set of additional variables considered to be significant. The estimating equation is of the form:

$$S_Y = \beta_0 + \beta_1 \ln FS + \beta_2 \ln Y + \beta_3 C_1 + , \ldots , + \beta_k C_{k-2} + \epsilon \qquad (1)$$

where:

S_Y = satisfaction with income, recoded to represent a zero to one welfare continuum,

FS = family size,

Y = monthly family income gross of tax,

C = a set of (k - 2) control variables,

β_0 thru β_k are the regression coefficients and ϵ is the disturbance term assumed to have the required characteristics. The control variables in the present context are entered as dummy variables and include, in this initial specification (note variable abbreviations in parentheses):[2]

* Financial situation in 1979 judged better than in 1978 (TB),
* Financial situation in 1979 judged worse than in 1978 (TW),
* Financial situation in 1980 expected to be better than in 1979 (NB),
* Financial situation in 1980 expected to be worse than in 1979 (NW), and
* Family reference person is age 65 or over (AGE).

Depending on theoretical considerations and the reasonableness of the results these variables may be directly included in the scale, for example, differentiating the scale by age of head, or may be used to control for effects that one does not want the scale to reflect. More information about the satisfaction measure and the income and the control variables is available from the author on request.[1]

Estimation results. Based on data from the fifth and sixth waves of the 1979 Research Panel of the Income Survey Development Program (ISDP) (see Ycas and Lininger 1981), (1) is estimated as

$$\begin{aligned} S_Y = {} & \underset{(6.24)}{0.140} - \underset{(-5.01)}{0.0224 \ln FS} + \underset{(19.59)}{0.0628 \ln Y} + \underset{(11.08)}{0.0715 TB} \\ & - \underset{(21.01)}{0.1139 TW} - \underset{(-4.23)}{0.0276 NB} - \underset{(-9.53)}{0.0508 NW} + \underset{(11.83)}{0.0763 AGE} \end{aligned}$$

$(R^2 = .263, N = 5{,}067)$

This rather simple model explains a little more than 25 percent of the variation in satisfaction with income. All variables in the model are statistically significant at conventional levels (t-values are in parentheses). As expected, income is positively related to satisfaction with income and the effect of family size is negative; holding income constant, satisfaction with income decreases as family size increases. The age coefficient has the expected sign and indicates that, controlling for the other variables in the model, the aged require less income to reach a given level of satisfaction than the non-aged. However, as we shall see, the magnitude of the age effect is much larger than seems reasonable.

Other things being equal, individuals who felt that their financial situation in 1979 was better than 1978 are more satisfied than others, while those who felt that their situation in 1979 was worse than in 1978 derive a great deal less satisfaction from a given level of income. Those who expected their financial situation to change in 1980, whether for better or worse, are considerably less satisfied than others at similar income levels. In fact, respondents'subjective assessments of financial change (retrospective for 1979 compared to 1978 and prospective for 1980 compared to 1979) are more important in accounting for variation in satisfaction with income than current income and family size.[3]

Deriving the scales. The purpose of an equivalence scale is to establish the relative incomes necessary to provide equal "satisfaction", "welfare", or "utility" for families in differing situations. In the present context we wish to construct a scale taking into account family size and age of head and will use non-aged families of size four as the reference family.

Restricting ourselves for the moment to simply family size effects among the non-aged, we need to know the ratios of income

$$\frac{Y_j}{Y_4} \qquad j = 1,2,3,5,6,\ldots,n$$

such that satisfaction with income (Sy) is the same for families of size one to three and five or more as for families of size four, that is:

$$Sy_j = Sy_4 \qquad (2).$$

Using equation (1), less the term representing the set of control variables which will be reintroduced below (see (7)),

$$Sy_j = \beta_0 + \beta_1 \ln FS_j + \beta_2 \ln Y_j + \epsilon \qquad (3)$$

and

$$Sy_4 = \beta_0 + \beta_1 \ln FS_4 + \beta_2 \ln Y_4 + \epsilon \qquad (4).$$

Combining (3) and (4) as suggested by (2), simplifying and arranging the terms for income and family size on opposing sides, and dividing through by β_2 gives:

$$\ln Y_j - \ln Y_4 = \frac{\beta_1(\ln FS_4 - \ln FS_j)}{\beta_2} \qquad (5).$$

Finally, taking the antilog of each side gives

$$\frac{Y_j}{Y_4} = \frac{FS_4}{FS_j}^{\left(\frac{\beta_1}{\beta_2}\right)} \qquad (6)$$

or the ratios of income necessary to leave families of size "j" as satisfied with their incomes as are families of size four.

Introduction of an additional variable or variables can be accomplished by inserting the terms $\beta_3 C_1 + \ldots, + \beta_k C_{k-2}$ from (1) in (3). The result, after the simplification and rearrangement represented by steps (5) and (6), is

$$\frac{Y_j}{Y_4} = \frac{\dfrac{FS_4}{FS_j}^{\left(\frac{\beta_1}{\beta_2}\right)}}{\exp\left[\dfrac{\beta_3 C_1}{\beta_2}\right] \times \ldots \times \exp\left[\dfrac{\beta_k C_{k-2}}{\beta_2}\right]}$$

the equivalence ratio net of the effect of a given control variable(s) divided by the effect of that variable (the exponentiated product of the variable and its coefficient divided by the coefficient of income).

The equivalence scales for non-aged and aged families, taken directly from equations (6) and (7), are shown below. Note that the values for the aged are too low to be given much credence. This point will be elaborated on further in the later sections of the paper.

Family size	Age of reference person	
	Under 65	65 or over
One....................	61	18
Two....................	78	23
Three..................	90	27
Four...................	100	30
Five...................	108	--
Six....................	116	--
Seven..................	122	--
Eight..................	128	--
Nine...................	134	--
Nine or more[a]	143	--

(--) - Not calculated.
[a]Based on an average family size of 10.9 for the nine and over group.

Results for the non-aged scale compared with other approaches.

Studies using direct subjective measures.-- Eight scales estimated using direct subjective measures are shown in table 1, together with the scale first presented in this paper. Six different data sets are involved, and all three types of direct subjective measures are represented: living level measures (6 scales); income metric measures (1 scale), and current income rating measures (2 scales).[4]

Of course in some sense the scales are not fully comparable. Important differences in study design and implementation cannot be controlled for. Thus, it would be unrealistic to expect complete uniformity in the various scales. However, there are important commonalities among the studies. For example, all scales but [7], for which age effects are controlled, pertain specifically to units headed by an individual under age 65, and all but [8], which employed family-size dummies, used a natural log specification for unit size.

The key attribute of these scales is their steepness. The formal way to assess their "steepness" is to compare their income-unit size elasticities, that is, the percentage change in income required to compensate for a given percentage difference in family size.[5] The elasticities for all nine of the scales are given at the bottom of table 1. Despite all the differences among the studies, there does seem to be considerable uniformity among the scales, as the elasticities for eight of the nine range in a relatively narrow band of .31 to .39. The elasticity for the "steepest" scale is .47, indicating that it is on the order of 20 - 50 percent "steeper" than the others.[6]

To give the flavor of what these income-unit size elasticities mean in practical terms, the eight "flatter" scales indicate that roughly 25 - 30 percent more income is required to maintain equivalent well-being as family size doubles (as

Table 1.--Selected subjective unit size equivalence scales for nonaged units by type of subjective measure and data source

Unit Size	Living Level Scales						Income Metric (WFI)	Income Satisfaction	
	Boston Social Standards Survey	Dubnoff & Strate		Minimum Income					
		Poor	Getting Along	Wisconsin (BNS)	1979 ISDP Wave 6	Non-Monetary Work Rewards Survey	Wisconsin (BNS)	1972 SRC	1979 ISDP Wave 6
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
One............	64[a]	64	58	52	65	62	60	58	61
Two............	80	80	76	72	81	78	78	87	78
Three..........	91	91	89	87	91	90	90	91	90
Four...........	100	100	100	100	100	100	100	100	100
Five...........	107	108	109	111	107	108	109	108	108
Six............	114	114	117	121	113	115[a]	116	--[b]	116
Seven..........	119	120	124	130	119	121[a]	123	--[b]	122
Eight..........	124	125	130	139	124	127[a]	129	--[b]	128
Unit size elasticity[c]...	.32[d]	.32	.39	.47	.31	.35[d]	.37	.37	.36

(--) - Not available.

[a]Derived using an unrounded elasticity as estimated from the published scale values. See also note (c) below.
[b]Not calculated.
[c]Where elasticity is estimated as "b" from the regression lnYR = a + blnUS with "YR" representing scale value (income ratio) and "US" representing unit size. Note also that this is equivalent to using, when available, the coefficients for unit size and income as estimated in each of the studies.
[d]Calculated from the published scale values. See note (c) above.

from 2 to 4 or 4 to 8 persons). For the BNS minimum income scale, about forty percent more income is required to yield an equivalent level of welfare across such family size differences.

In sum, this preliminary version of an ISDP income satisfaction scale appears to be quite similar to other "subjective" scales estimated to date for the U.S.[7] There also appears to be at least some regularity amomg the scales estimated so far on the basis of direct subjective assessments. How do these "subjective" scales compare to those derived using traditional approaches?

Comparisons with scales based on dietary need and consumer behavior.-- Seven different "objective" scales are shown in table 2. Three widely used scales are shown in the first three columns of the table. The Thrifty Food Plan, developed by the USDA (col. 2), is a dietary adequacy scale and represents the ratio of food plan costs necessary to maintain equal nutritional adequacy by family size (Kerr and Peterkin 1976; Peterkin 1976). The Federal Poverty Line scale (col. 1) is largely based on an earlier version of the USDA Thrifty Food Plan, the Economy Food Plan (Orshansky 1965). As reflected in the poverty line scale values for units size five and above, it was slightly less steep than the Thrifty Plan. However, the most pronounced differences between the food plan scale and the poverty scale are for two, and more especially one-person, units. As noted earlier these differences stem from intuitive adjustments made to the poverty scale in order to account for supposed "diseconomies" experienced by small units in the consumption of nonfood items. The BLS scale (col. 3) was derived by estimating the total expenditure levels at which the percent of total expenditures devoted to food remained constant by family size, i.e. it is a classic Engel type scale (Jackson 1968). Interestingly, it is very similar to the dietary adequacy scale up through families of size five.

The Iso-prop scales shown in columns 4 and 5 are also Engel scales, i.e., they were estimated by holding constant expenditures for food or necessities as a percentage of total consumption expenditures for families of different size (Watts 1967). The Iso-prop food scale is quite similar to the Thrifty Food Plan and BLS Family Budget scales. The necessities scale differs from the other four scales in that it is noticeably less steep.

The food-based scales indicate that a family of eight would need about seventy percent more income than a four-person family to be equally well-off. The broader based necessities scale suggests that the family of eight would need only about forty percent more. At family sizes below four, of course, the necessities scale is appreciably more compressed than the three food based scales. Because of the ad hoc adjustments incorporated into the poverty line scale for units size one and two, at the lower end the Iso-prop necessities scale and the poverty scale are roughly similar to each other, suggesting, perhaps, the essential reasonableness of those adjustments.

The final two scales, estimated using the extended linear expenditure system (ELES) on the basis of all categories of current consumption expenditures, are flatter even than the iso-prop necessities scale (van der Gaag and Smolensky 1982, Danziger, van der Gaag, Taussig and Smolensky 1984). The contrast between the ELES scales and the food based scales is, quite naturally, even more marked.

The comparison, then, of these different types of

Table 2.--Selected non-aged equivalence scales estimated using dietary standards, food share, or more broadly based definitions of consumption

Unit Size	Federal Poverty Line [1]	Thrifty Food Plan [2]	BLS Family Budget[b] [3]	Iso-prop Index[a] Food [4]	Iso-prop Index[a] Necessities [5]	ELES One[b] [6]	ELES Two[b] [7]
One............	50	33[b]	39	36	51	57	59
Two............	65	59	61	60	72	80	80
Three..........	78	80	83	81	87	92	91
Four...........	100	100	100	100	100	100	100
Five...........	118	120	124	118	111	110	111
Six............	134	144	148[c]	135	122	117	117
Seven..........	152	161	--[d]	152	131	--[d]	--[d]
Eight..........	170	174	--[d]	168	140	--[d]	--[d]
Income-unit size Elasticity[e]...	.60	.80	.71[f]	.74	.48	.39	.38

[a]Scale is for poverty level incomes and is net of age effects rather than for the non-aged per se.
[b]Weighted using ISDP population counts.
[c]Six or more.
[d]Not available.
[e]See note (c) table 1.
[f]Estimated using scale values for units of size 1-5 only.

Table 3.--Selected aspects of aged equivalence by type of scale and data source

Size of unit and sex of head	Objective Measures				Subjective Measures					
	Federal Poverty Line	BLS Family Budget	ELES		Minimum Income		Dubnoff & Strate		Income Satisfaction	
			One	Two	Wisconsin BNS	1979 ISDP Wave 6	Poor	Getting Along	1972 SRC	1979 ISDP Wave 6
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
Part I: Aged as percent of non-aged equivalent incomes for one and two-person units (non-aged unit of size 2 = 100)										
One[a]..........	71	51	35	51	82	40	58	57	34	23
Male.........	72[b]	--	51	62	78	52	--	--	--	--
Female.......	71[b]	--	31	49	83	36	--	--	--	--
Two...........	90[c]	93	73[c]	84[c]	107[c]	64[c]	74	75	50	30
Part II: Aged one-person as a percent of aged two-person equivalent incomes										
Both sexes[c]...	79	55	48	61	77	63	78	76	72	78
Male.........	80[b]	--	70	74	73	81	--	--	--	--
Females......	79[b]	--	42	58	78	56	--	--	--	--

(--) - Not applicable.
[a]Weighted average for scales including a distinction by sex of head.
[b]As existed prior to 1981. Since 1981 distinctions by sex have been eliminated.
[c]Husband-wife couple or male head.

objectively based scales suggests that as one moves away from scales based on food (whether nominally grounded in dietary adequacy or an Engel technique) to scales based on more inclusive categories of consumption, family equivalence scales tend to flatten considerably; perhaps to a degree, represented by the ELES scales, that some observers would find hard to accept (Deaton and Muellbauer 1980, pp. 201-206). Nonetheless, the pattern presented is rather marked and unambiguous.

Turning now to a comparison of the direct subjective scales to this set of "objective" scales, it is immediately obvious that the subjective scales bear little resemblance to the food-based scales shown in table 2. In fact, with the exception of the BNS minimum income scale, the subjective scales look quite similar to the ELES scales, the flattest in the objective category. The BNS scale, on the other hand, is virtually indistinguishable from the Iso-prop necessities scale. However, since there doesn't seem to be any readily identifiable substantive basis for its greater "steepness", this doesn't seem to be especially noteworthy.[*]

Equivalence and the aged

One of the main issues which we will be addressing in this project concerns the economic needs of the aged both in comparison to the non-aged and among major subgroups of the aged population itself.

Table 3 contains information from ten different scales which distinguish units with heads age 65 or older. Four of the scales were derived using objective measures (columns 1-4) and six, including estimates from the sixth wave ISDP income satisfaction item, are based on direct subjective measures (columns 5-10).

There are three important questions in the area of aged equivalence which are addressed by information in the table: 1) are the needs of the aged systematically different from those of the non-aged? 2) How do the needs of aged one and two-person units compare? 3) when living alone, i.e., in single-person units, are the needs of aged men and women different?

The table is divided into two parts. The first three rows of the table deal with the relative needs of one and two-person units and males and females indexed to the needs of a non-aged unit of size 2 = 100. Rows 2 and 3 show the equivalence ratios for males and females living alone expressed as a percentage of a non-aged two-person unit. Row 1 gives the average for single-person aged units weighted on the basis of the numbers of aged men and women living alone in the noninstitutional population. The information in rows one to three is then re-expressed in part II of the table (the last three rows) using aged units of size two = 100 as a base. Finally, since the base for the percentages in part one of the table is non-aged units of size two = 100, the row for units of size two (row 4) isolates the overall "age" effect.[*]

Turning to the first issue, what do these scales imply about the needs of the aged compared to the non-aged? It is often asserted that the aged require less than the non-aged to sustain a comparable level of living (see, for example Danziger, van der Gaag, Smolensky and Taussig 1984). They are able to maintain equivalent nutritional status at lower levels of food consumption, they often live in fully paid for owner-occupied housing, no longer experience the expenses associated with working and, if comparisons are made on the basis of before-tax income, generally a smaller portion of the income of the aged is devoted to taxes. On the other hand, they may experience greater out-of-pocket health expenses, and concerns about economic security could produce a shift in the focus of economic preoccupations from consumption to savings. What do these ten equivalence scales indicate regarding the relative needs of the aged and non-aged?

Indeed, all but the BNS minimum income scale indicate that controlling for family size, the aged require less than the non-aged to attain an equivalent level of economic well-being. Since the standard error for the age coefficient underlying the BNS scale is virtually as large as the coefficient (Colasanto, Kapteyn and van der Gaag 1984, p. 134), no substantive importance need be attached to its failure to conform to the general picture of relatively lower material needs for aged families. Apart from this rather commonplace finding that the aged's "needs" are less than those of the non-aged, the diversity in the estimates is rather striking. It is probably noteworthy that the two "food-based" measures yield the highest aged/non-aged ratios (about 90). The two variants of the ELES scale, while lower than the food based scales, differ somewhat from each other. The version denoted as ELES I (van der Gaag and Smolensky 1982) indicates the aged require only 73 percent as much as the non-aged to be equally well-off, while ELES II (Danziger, van der Gaag, Smolensky and Taussig 1984) indicates that the aged's needs are 84 percent of those of the non-aged.

With the exception of the BNS scale, which can be discounted because of the unreliability of the age dimension of the scale, the direct subjective scales are all at or below 75 for the aged. Three of the five are in the 64 to 75 range, while the other two, the 1972 SRC and 1979 ISDP income satisfaction scales, are considerably lower (50 and 27 respectively). In fact, the ISDP income satisfaction scale is too low to be taken seriously.

Thus it would seem that the wide range presented by the estimates precludes any useful general statement about the relative needs of the aged and non-aged based on these studies. Apparently a good deal more care is needed to lay out the nature of the problem, to insure that variables are defined in a theoretically relevant manner and to take care that the models are correctly specified.

The two remaining issues addressed in table 3, the possible impact of gender on needs and the comparative needs of one and two-person aged units, are closely related and central to aged size equivalence issues. This is because the vast majority of aged individuals live alone or with just one other person and approximately four-fifths of aged individuals living alone are women.

Gender effects.--Five of the ten scales provide distinctions by sex of head as well as age. Four indicate that the needs of aged women living alone are less than those of aged men. In these four studies, estimates of women's needs vary from only very slightly less than men's (viz. the U.S. poverty thresholds as constituted prior to 1981) to substantially less (the two ELES scales and the ISDP minimum income scale).

While the Wisconsin BNS minimum income scale indicates that women's needs are greater, the standard error for the sex of head coefficient is very large in comparison with the coefficient

(Colasanto, Kapteyn and van der Gaag 1984, p. 134). Some limited experimentation with universe and variable specifications with the ISDP data have produced coefficients suggesting greater needs for female heads also, and in some cases the effect was even statistically significant. However, given the universe definition employed in estimating the present version of the scale, sex of head was not statistically significant and, as noted earlier, was dropped from the model.

Thus, notwithstanding the large gender effects shown in some of the scales, it seems we are some way from being able to demonstrate a consistent sex effect, much less reliably gauge its magnitude. And even if its presence could be firmly established, its precise significance would be far from clear because of the very low incomes of many aged women.

One versus two-person units.-- Given the problematic nature of gender effects, the question of one-person/two-person unit equivalence among the aged is necessarily somewhat clouded. The one to two-person ratios for the ten scales are distributed over a considerable range (see row 5 of the table). Six of the scales indicate that the needs of an aged person living alone are between 70 to 79 percent of an aged couple's needs. Indeed, five of the ten scales yield one-person/two-person ratios in the range of 76-79. The value for the Federal poverty line (79) places it in this group of five. However, the ELES I and BLS scales suggest a nearly per capita relationship (48 and 55). The ELES II and ISDP minimum income scales are also noticeably below the group of six (61 and 63 respectively).

The sex effect is evident especially in the two ELES scales and the ISDP minimum income scale and would appear to be largely responsible for the fact that the overall one-person/two-person ratios for these scales are so low. Again, a good deal more work is needed to understand the basis of such widely varying estimates.

Some elaborations of the initial model.

Further exploration of the age effect and the specification of the family size variables is obviously called for. While in no sense definitive, a minimal effort was possible after return from the meetings. Some interesting results were obtained and will be reported on.

Four basic modifications were made to the naive model that we estimated initially: 1) an interaction term was added for family size and income, 2) a dummy variable was introduced for one-person aged units, 3) a substantially expanded set of control variables thought to be related to aged/non-aged differences was added, and 4) a crude approximation of after-tax income was substituted for the before-tax variable.

The family size/income interaction term was added principally to test if family size effects vary by income level. This possibility has been commonly noted and has received some empirical support (e.g. Watts 1967, Seneca and Taussig 1971, and Deaton 1982). Both theory and empirical evidence suggest that the addition of family members in better off families requires a smaller proportional increase in expenditures (e.g., Deaton 1982, p. 43). Thus, we might expect a size equivalence scale for low income families to be "steeper" than a scale for families with incomes in the middle or upper part of the distribution.

The other modifications were made principally to find out something more about the nature of the age effect. The dummy variable for aged single-person units was added because aged singles confront very different conditions from aged couples. They tend to be older and have notably lower incomes. If their reduced circumstances were to be systematically related to lowered aspiration levels, specifications assuming a constant age effect by unit size could prove distortive and might, in part, account for the excessively large age effect estimated from the original simple model. The remaining modifications were undertaken in order to see to what extent the age effect in the initial model could be explained by taking into account some of the factors which might contribute to lower before-tax money income needs for the aged. Given the variables at hand on the analysis file, home tenure, number of earners, value of durables, and amount of other fungible assets (excluding own home) could be added to the initial model. As a matter of general interest, a dummy for work disability among the non-aged was also incorporated. Finally, a few runs were made to experiment with the impact of using after-tax as opposed to before-tax income as a predictor variable. The results are given in table 4.[10]

Results for the non-aged.--The interaction term was included in two extensions of the naive model . . . as a sole addition to the initial variable set (Model II), and together with the dummy for aged units of size 1 (Model IV). The result is given in table 4 (Models II and IV). The coefficient for the interaction term has the expected sign in each instance. It is positive, indicating that the lower the income level, the higher the elasticity of income with respect to family size and the steeper the resulting equivalence scale. However, the magnitude of the coefficient is affected by the other variables which were added to explore the age effect. For example, when the dummy variable for aged units of size one and the family size interaction term are both present (Model IV), the interaction effect is halved and the t-value drops from the margin of significance at the .05 level (single-tailed test) to well below conventional levels of significance (Model IV compared to Model II). This would seem to indicate that a good deal of the gross interaction affect is attributable to the special circumstances of one-person aged units.

The impact of the interaction term on equivalence scales per se for the non-aged is shown in table 5. Coefficients from Model II (the interaction term only added to the initial set of variables) and Model IV (the interaction term and the dummy for aged units of size one both added to the initial variable set) were used to construct scales at four different reference income levels keyed to four-person families (the average income of poor families, the weighted poverty threshold income, median income and twice the median, all in terms of four-person family incomes).[11] The scales derived from the initial model and Model III (the initial model plus the term for aged units of size one) are included in the table for purposes of comparison.

Clearly the elasticity of income with respect to family size varies noticeably by income level. The elasticity estimated from the initial model was .36. Using Model II, the average elasticity ranges between .49 for families living in poverty to .19 for those living at satisfaction levels experienced by four-person families with incomes at twice the median for four-person families.

Estimates based on Model IV also reveal important

Table 4.--Coefficients for the initial model and selected alternative specifications

Selected coefficients[a]	Model						
	Initial	II[b]	III[c]	IV[d]	V[e]	VI[f]	VII[g]
Intercept...........	.1401	.1825	.1278	.1499	.1954	.1734	.1583
	(6.24)	(5.27)	(5.39)	(4.06)	(7.70)	(6.71)	(5.69)
lnFS(β_1)............	-.0224	-.0791	-.0189	-.0474	-.0230	-.0283	-.0252
	(-5.01)	(-2.03)	(-4.08)	(-1.26)	(-4.77)	(-6.04)	(-5.23)
lnY(β_2).............	.0628	.0565	.0640	.0608	.0481	.0578	.0596
	(19.59)	(11.71)	(19.80)	(11.39)	(12.07)	(15.08)	(15.34)
lnFS*lnY(δ).........	...	.0082	...	.0041	...	...	...
		(1.61)		(0.76)			
Perceived financial change, (1979 vs. 1978):							
Better(β_3)........	.0715	.0715	.0711	.0711	.0752	.0755	.0751
	(11.08)	(11.08)	(11.00)	(11.01)	(11.85)	(11.63)	(11.58)
Worse(β_4).........	-.1139	-.1138	-.1137	-.1136	-.1116	-.1141	-.1138
	(-21.01)	(-21.00)	(-20.96)	(-20.96)	(-20.92)	(-20.88)	(-20.83)
Expected financial change, (1980 vs. 1979):							
Better(β_5)........	-.0286	-.0291	-.0283	-.0286	-.0231	-.0253	-.0249
	(-4.23)	(-4.30)	(-4.18)	(-4.21)	(-3.48)	(-3.70)	(-3.65)
Worse(β_6).........	-.0508	-.0507	-.0508	-.0508	-.0522	-.0499	-.0498
	(-9.53)	(-9.52)	(-9.53)	(-9.52)	(-9.98)	(-9.29)	(-9.29)
Aged(β_7)............	.0763	.0751	.0642	.0646	.0416	.0722	.0604
	(11.83)	(11.59)	(8.35)	(8.38)	(2.84)	(11.12)	(7.77)
Aged, 1-person unit(β_8)...........	...	...	.0320	.0293	.0278	...	.0309
			(2.90)	(2.53)	(2.39)		(2.76)
R^2..................	.263	.263	.264	.264	.297	.253	.254
Number of Cases = 5,067							

Note: t-values in parenthesis.

[a]Coefficients for the additional control variables incorporated in Model V not included in the table. See text and text footnote 10.
[b]Initial model plus interaction term for income and family size.
[c]Initial model plus a dummy variable for aged single-person units.
[d]Initial model plus interaction term for income and family size and a dummy variable for aged single-person units.
[e]Initial model plus a dummy variable for aged single-person units, plus the additional set of control variables.
[f]Initial model substituting after-tax for before-tax income.
[g]Initial model substituting after-tax for before-tax income plus a dummy variable for aged single-person units.

differences in elasticity by income level. While the interaction coefficient is not significant statistically, it does produce marked differences in the scale. Comparing the extremes, for example, the family size elasticity for poor families is about 1.6 times higher than for families at the highest level.

In sum, the data set does offer some support for the notion of interaction between income and family size and illustrates that the effect could be of considerable practical significance. In fact, the contrast noted earlier in steepness between food-based scales and those estimated on the basis of broader consumption sets might be attributable, in part, to the failure to appropriately account for such interaction effects. However, a great deal of the interaction seems to arise from the special circumstances of aged one-person units. After removing this effect, the interaction term still yields scales with noticeably different elasticities by income level, but the interaction term is not statistically significant. Thus, in future work it would seem advisable to explicitly test for interaction effects of this type. Other more sophisticated ways of pursuing possible interaction effects are available as well, and should be considered (see for example, Deaton 1982, p. 41).

Impact on estimates for the aged.-- The most consistent and important impact on age effects and, consequently, aged equivalent income, is attributable to the addition of the dummy variable for aged units of size one (see tables 4 and 6). The sign is in the expected direction, and the coefficient is statistically significant in each of the four models which include the variable (III, IV, V and VII). When the dummy is present, it leads to three results. The estimate for the aged equivalent income is lower for one-person than two-person units. One-person unit equivalent income is also lower than estimated on the basis of the initial model. Finally, the ratio for two-person units increases somewhat over the initial model, but not dramatically (from about 30 to about 36 to 37 when scaled to non-aged incomes = 100). These results suggest a difference in the income - needs relationship (needs, of course, as measured by income satisfaction) between one and two-person aged units, but not one large enough to explain away the very large age effect uncovered in the initial model.

The attempt to net out the effects of work activity and assets (including owner-occupied housing) had some slight additional impact. All additional variables were of the expected sign, i.e., the more assets the less income required to reach a given level of income satisfaction; earnings activity and nonwork due to disability among the non-aged were associated with less satisfaction at a given level of before-tax income.[12] However, the combined effect of the additions was not marked and was most evident for two-person units where aged equivalent income increased from 37 (Model III) to 42 (Model V).

Allowing for family size/income interaction had slight, if any, effect when taken alone (Model II), and had no visible impact when present in conjunction with the dummy for aged units of size one (Model IV vs. Model III). Use of the approximation of after-tax income also had very little net effect (Initial model vs. Model VI and Model III vs. Model VII) but the results were somewhat mixed. The coefficient for the age 65 plus dummy (β_1) did decline slightly (by about 5-6 percent), but the coefficient for income (β_2) dropped somewhat more (about 10-11 percent). As a result, and contrary to expectations, aged equivalent income on an after-tax basis is essentially the same as on a before-tax basis. Based on the supposition that after-tax income would be more closely tied to satisfaction than before-tax income, one would have supposed that the income coefficient would be larger on an after-tax as opposed to before-tax basis. This contrary finding is somewhat puzzling and may arise from the somewhat crude simulation of after-tax income employed.

The main point raised earlier in the discussion of table 3 is little affected. The aged/non-aged equivalence ratios stemming from the sixth wave ISDP income satisfaction measure remain the lowest of any considered and do not appear to be credible. Almost certainly an explanation, if one is to be found, lies along different avenues than those touched on here.

At the same time, and not surprisingly, the dummy variable for aged units of size one did have a marked effect on the one-person/two-person equivalence ratios for the aged. Whereas the initial model yielded a one vs. two-person equivalence ratio of 78, the models including the extra dummy for aged one-person units (again Models III, IV, V and VII) yielded ratios of no more than 50. Of the ten scales considered earlier, only the one denoted ELES I yielded such a low value.

Summary and conclusions.

This paper has provided a general discussion of the use of direct subjective assessments of income in equivalence scale applications and has presented a heuristic empirical example of the estimation of an equivalence scale using such measures.

The scale estimated for this paper was compared with scales obtained by others using direct subjective measures and more traditional approaches. The comparisons focused on four aspects of equivalence: 1) the impact of family size on needs among non-aged families, 2) the relative needs of aged and non-aged families, 3) gender effects among the aged, and 4) the needs of one and two-person aged units.

The family size issue was reviewed in terms of the steepness of the scales. As a group, the scales based on direct subjective measures appeared to be noticeably less steep than scales based on dietary needs or food consumption and rather similar to scales based on more broadly defined consumption sets such as necessities and especially total current consumption, i.e., the two ELES scales. However, some limited experimentation with alternative specifications of the model estimated in this paper suggested that this contrast may be due in part to a failure to take proper or full account of possible interaction between family size and income.

Review of the findings concerning equivalence issues for the aged was inconclusive. While there was a clear tendency for the needs of the aged, as defined in these studies, to be less than the needs of the non-aged, estimates of the magnitude of the needs differential varied widely. No consistent pattern was discernible in regard to the other two issues central to question of aged equivalence (gender effects and the needs of one vs. two-person aged units).

It would seem that before a reasonably orderly set of empirical results can be expected to emerge for the aged much more work will have to be done to

Table 5.--Non-aged unit size equivalence scales based on selected alternative specifications of the initial model

Model[a]	Unit Size								
	1	2	3	4	5	6	7	8	IFSE[b]
Initial model......	61	78	90	100	108	116	122	128	.36
Model II									
Income level									
AP($392/m).......	48	71	88	100	110	119	126	133	.49[c]
PT($616/m).......	52	74	89	100	109	116	123	128	.43[c]
M($1,882/m)......	65	82	93	100	106	110	114	118	.28[c]
2M($3,763/m).....	77	88	95	100	104	107	109	112	.19[c]
Model III..........	66	81	92	100	107	113	118	123	.30
Model IV									
Income level									
AP($392/m).......	59	78	90	100	108	115	121	126	.36[c]
PT($616/m).......	62	79	91	100	107	113	119	124	.33[c]
M($1,882/m)......	68	83	93	100	106	110	115	118	.26[c]
2M($3,763/m).....	73	86	94	100	105	109	111	115	.22[c]

INCOME LEVEL KEY:
AP - 1/12 the average annual income of poor 4-person families, 1979.
PT - 1/12 the annual poverty threshold for 4-person families, 1979.
M - One-twelfth the median income of 4-person families, 1979.
2M - One-twelfth twice the median income of 4-person families, 1979.

[a]For description of models see notes to table 4 and text.
[b]Income-family size elasticity.
[c]Average elasticity. Elasticity for the lower part of the scale is higher than for the upper part of the scale.

Table 6.--Selected aspects of aged equivalence using the initial model and selected extensions

Unit size	Model[a]						
	Initial	II	III	IV	V	VI	VII
	Part I: Aged as percent of non-aged equivalent incomes for units of size one and two						
One-person.....	30	26	22	21	24	29	22
Two-person.....	30	30	37	36	42	29	36
	Part II: Aged one-person as a percent of aged two-person equivalent incomes						
Aged units of size 2 = 100..	78	70[b]	50	47[b]	39	71	44

[a]See notes to table 4 and text for description of models.
[b]At a welfare level equivalent to that of four-person families with incomes at the median for four person families.

refine existing independent variables and introduce new ones and to develop more appropriate models. Likewise, considerable additional work remains to be done in order to develop the equivalence scale presented here beyond the frankly preliminary stage.

Footnotes.

*Dan Radner, Ben Bridges, Clarise Lancaster, Arie Kapteyn and Steve Dubnoff provided a number of very helpful suggestions and criticisms. My thanks to Sharon Johnson for working through the complicated fifth and sixth wave ISDP files to produce the matched analysis extract. Ray Becker and Mike Rozdilski also assisted with the data processing, and Weltha Logan helped with editorial details.

†The algorithm used to simulate after-tax income for the 1984 Proceedings version of this paper contained keying errors which had a substantial affect on some of the parameter values (namely the intercept and the coefficients for family size and income) for the models which incorporate after-tax income. Fortunately, the errors tended to offset each other and so their net impact on aged/nonaged equivalence was slight. These errors have been corrected in the current version of the paper.

[1]The author encourages interested parties to contact him at the following address: Social Security Administration, 1875 Connecticut Ave., N.W., Rm. 320-N, Washington, D.C. 20009

[2]An alternative specification included female head as a control variable. The sign was negative, but not significant ($t = -0.81$). The model was reestimated with the five control variables listed.

[3]Based on comparison of the sums of squares uniquely associated with the model variables (as gauged by successively adding each variable as the last in the model), the sums of squares due to current income and family size amount to only about 60 percent of the sum of squares associated with the four subjective change variables. The powerful effect of perceived financial change on current satisfaction with income, especially change for the worse, has been documented earlier by Vaughan and Lancaster (1980) and Dubnoff, Vaughan and Lancaster (1981).

[4]Sources for the scales in table 1 are as follows: [1] Rainwater (1974), tables 5-2 and 5-4, pp. 102,105; [2] & [3] Dubnoff and Strate (1984), table 2, p. 18; [4] & [7] Colasanto, Kapteyn and van der Gaag 1984, pp. 127-138; [5] Danziger, van der Gaag, Taussig and Smolensky (1984), table 2, p. 53; [6] Dubnoff (1982), table 1, p. 10; [8] Dubnoff, Vaughan and Lancaster (1981), table 2, p. 351; [9] This paper.

[5]The exponent on the right hand side of equation (6) is in fact the negative elasticity. The elasticities appearing at the foot of tables 1 and 2 function in the same fashion, but because they are positive, they operate on the ratio FS_j/FS_4 instead of FS_4/FS_j as in (6).

[6]While the labeling of the BNS minimum income scale as "steeper" than the others is somewhat arbitrary, its elasticity is more than two standard deviations above the mean elasticity of the nine scales presented in the table.

[7]The fact that this estimate is preliminary is to be stressed. Limited experimentation with alternative models and variable definitions, some of which is reported on in the last section of the paper, indicates that a range of elasticities may be obtained. At this early stage of our research we have not developed a clear basis for selecting one of these alternatives over another.

[8]In light of the findings reported in the last section of the paper, perhaps this statement is too unequivocal. Unlike most of the other scales, the BNS minimum income scale pertains to a poverty level income, and my own analysis yields an average elasticity of .43 for a poverty threshold welfare level when the dummy for aged units of size one is left out of the model (the model for the BNS scale contained no such dummy). Clearly, the Dubnoff-Strate results indicate that subjective scales estimated for poverty income welfare levels are not necessarily "steeper". One could also protest that the ISDP minimum income scale is not "steeper" than the others, but it is also true that the income levels associated with the scale are very close to the median rather than at the poverty level. In any case, the difference between the BNS minimum income scale and the BNS WFI scale, essentially estimated at median income, has also been encountered in Dutch studies (Goedhart, et al., 1977), although the difference is not as pronounced. Until recently, members of the original Leyden group had not made much of the difference, but recently some attention has been given to this matter (personal communication, Arie Kapteyn).

[9]Strictly speaking this statement is only true for scales in which age effects do not vary by unit size. This condition holds for all scales estimated with a dummy specification for age, as is the case for all but two of the scales. While the poverty line age effect was not explicitly derived on the basis of a dummy specification, the aged/non-aged ratios for one and two-person units are the same (.90). However, for the BLS scale the aged/non-aged equivalence ratio for 1-person units is .72. as opposed to the ratio of .93 based on a comparison of two-person units as shown in the table.

[10]An after-tax specification of income is clearly the variable of choice for most applications, but the available variable assumes use of the standard deduction only for the Federal income tax and thus underestimates after-tax income levels for about the upper third of the distribution. As this project advances we hope to develop a better measure of after-tax income and employ it more generally in our analyses.

[11]The income (Y) for a family of size "j" equivalent to that of a given reference level income for a family of four (Y_4) may be derived from the **exponentiated value of the following expression:**

$$\frac{\dfrac{\beta_1(\ln FS_4 - \ln FS_j)}{\beta_2} + \ln Y_4\left(1 + \dfrac{\delta}{\beta_2}(\ln FS_4)\right)}{\left(1 + \dfrac{\delta}{\beta_2}(\ln FS_j)\right)}$$

where δ is the coefficient for the family size/income interaction term. Details on the derivation of this formula are available from the author.

[12]The t-values for all but one of the additional variables (living in means-tested public or publicly subsidized housing with a t-value of

1.41) range between 1.8 and 7.8. While the magnitude of a number of the coefficients seems rather large, this is probably attributable, at least in part, to the fact that the income definition employed in Model V is before-tax and so variables for homeownership and work presumably reflect tax effects in addition to whatever independent effects they may exert.

<u>References</u>.

Andrews, F.M. and S.B. Withey (1976).
"Social Indicators of Well-Being. New York:Plenum Press.

Colasanto, D., A. Kapteyn and J. van der Gaag (1984).
"Two Subjective Definitions of Poverty: Results from the Wisconsin Basic Needs Study." Journal of Human Resources, vol. 28. no. 1. pp. 127-138.

Danziger, S., J. van der Gaag, E. Smolensky and M. Taussig. (1984).
"Income Transfers and the Economic Status of the Elderly." appears as chapter 7 in M. Moon, ed. Economic Transfers in the United States. Chicago: University of Chicago Press.

Danziger, S., J. van der Gaag, M. Taussig and E. Smolensky. (1984).
"The Direct Measurement of Welfare Levels: How Much Does It Cost to Make Ends Meet?." Review of Economics and Statistics. vol. 66. no. 3, pp. 500-505.

Deaton, A. (1982).
"Inequality and Needs: Some Experimental Results from Sri Lanka." in Y. Ben-Porath, ed. Income Distribution and the Family. A supplement to volume 8. Population and Development Review, pp. 35-49

Deaton, A. and J. Muellbauer. (1980).
Economics and Consumer Behavior. Cambridge: Cambridge University Press.

Dubnoff, S.J. (1982).
"Validating Family Size Equivalence Scales." Boston: Center for Survey Research. University of Massachusetts. mimeo.

Dubnoff, S.J. and J.M. Strate (1984).
"How Much Income is Enough? Measuring the Income Adequacy of Retired Persons Using a Survey Based Approach." Boston: Center for Survey Research. University of Massachusetts.

Dubnoff, S.J., D.R. Vaughan and C. Lancaster. (1981). "Income Satisfaction Measures in Equivalence Scale Applications." in The 1981 Proceedings of the Business and Economic Statistics Section. American Statistical Association: Washington, D.C. pp. 348-352.

Goedhart, T., V. Halberstadt, A. Kapteyn and B.M.S. van Praag. (1977).
"The Poverty Line: Concept and Measurement." Human Resources. vol. 12, no. 4, pp. 503-520.

Jackson, C. (1968).
Revised equivalence scale for estimating income or budget costs by family type. Bulletin No. 1570-2. U.S. Department of Labor. Bureau of Labor Statistics.

Kakwani, N. (1977).
"On the Estimation of Consumer Unit Scales." The Review of Economics and Statistics. vol. 59. no. 4. pp. 507-510.

Kapteyn, A., and B.M.S. van Praag. (1976).
"A New Approach to the Construction of Family Equivalence Scales." European Economic Review. vol. 7. pp. 313-335.

Kerr, R. and B. Peterkin (1976).
"The Effect of Household Size on the Cost of Diets that are Nutritionally Equivalent." in The Measure of Poverty. Technical Paper XII. Food Plans for Poverty Measurement. Washington, D.C.: Department of Health and Human Services. pp. 64-87.

Muellbauer, J. (1980).
"The Estimation of the Prais-Houthakker Model of Equivalence Scales." Econometrica. vol. 48. no. 1. pp. 153-175.

Nicholson, J. (1976).
"Appraisal of Different Methods of Estimating Equivalence Scales and their Results." Review of Income and Wealth. series 22. no. 1. pp. 1-11.

Nicholoson, J. (1949).
"Variations in Working Class Family Expenditure." Journal of the Royal Statistical Society A. 112. pp. 359-411.

Orshansky, M. (1965).
"Counting the Poor: Another Look at the Poverty Profile." Social Security Bulletin. vol. 28. No. 1. pp. 3-29.

Peterkin B. (1976).
"USDA Food Plans, 1974." in The Measure of Poverty. Technical Paper XII. Food Plans for Poverty Measurement. Washington, D.C.: Department of Health and Human Services. pp. 3-31.

Pollak, R. and T. Wales. (1979).
"Welfare Comparisons and Equivalence Scales." American Economic Review. vol. 69. no. 2. pp. 216-221.

Prais, S.J. and H.S. Houthakker. (1955, 2nd edit. 1971).
The Analysis of Family Budgets. Cambridge: Cambridge University Press.

Rainwater, L. (1974).
What Money Buys: Inequality and the Social Meanings of Income. New York: Basic Books.

Rainwater, L. (1973).
"Poverty, Living Standards and Family Well-Being," Studies in Public Welfare. Paper No. 12. Sub-Committee on Fiscal Policy. Joint Economic Committee of the Congress.

Sen, A. (1982).
Choice, Welfare and Measurement. Cambridge, Mass.:The MIT Press.

Seneca, J. and M. Taussig. (1971).
"Family Equivalence Scales and Personal Income Tax Exemptions for Children." The Review of Economics and Statistics. vol. 53. no. 3. pp. 253-262.

Sydenstriker, E. and W. King (1921).
"The Measurement of the Relative Economic Status of Families." Quarterly Publication of the American Statistical Association. vol. 17. pp. 842-857.

United States Department of Agriculture. (1975).
"The Thrifty Food Plan." in The Measure of Poverty. Technical Paper XII. Food Plans for Poverty Measurement. Washington, D.C.: Department of Health and Human Services.

U.S. Department of Labor. Bureau of Labor Statistics. (1948).
Workers Budgets in the United States: City Families and Single Persons. BLS Bulletin 927.

van der Gaag, J. and E. Smolensky (1982).
"True Household Equivalence Scales and Charac-

teristics of the Poor in the United States." Review of Income and Wealth. series 28. no. 1. pp. 17-28.

van Herwaarden, F.G., A. Kapteyn and B.M.S. van Praag. (1977).
Twelve Thousand Individual Welfare Functions of Income." European Economic Review. vol. 9. pp. 283-300.

van Praag, B.M.S. (1971).
"The Welfare Function of Income in Belgium: An Empirical Investigation." European Economic Review. vol. 2. pp. 337-369.

van Praag, B.M.S. and A. Kapteyn. (1973).
Further Evidence on the Individual Welfare Function of Income: An Empirical Investigation in the Netherlands. European Economic Review. vol. 4. pp. 33-62.

Vaughan, D. and C. Lancaster. (1980).
"Applying a Cardinal Measurement Model to Normative Assessments of Income: Synopsis of a Preliminary Look." in The 1980 Proceedings of the Survey Research Methods Section. American Statistical Association: Washington, D.C.

Vaughan, D. and C. Lancaster. (1979).
"Income Levels and Their Impact on Two Subjective Measures of Well-being: Some Early Speculations from Work in Progress." in The 1979 Proceedings of the Social Statistics Section. American Statistical Association: Washington, D.C. pp. 271-276.

Watts, H.W. (1980).
"Special Panel Suggests Changes in BLS Family Budget Program." Monthly Labor Review. vol. 103. no. 12 (Dec.). pp. 3-10.

Watts, H.W. (1967).
"The Iso-Prop Index: An Approach to the Determination of Differential Poverty Income Thresholds." The Journal of Human Resources. vol. 2. no. 1. pp. 3-18.

Ycas, M. and C. Lininger. (1981).
"The Income Survey Development Program: Design Features and Initial Findings." Social Security Bulletin. vol. 44. no. 20, pp. 13-19.

COMMENT

Courtenay Slater, CEC Associates

The number and variety of SIPP papers being presented at these meetings -- in advance of actually having seen any data from SIPP -- certainly whet one's appetite for what is to come once data actually become available. The three papers on which I will be commenting raise two general points to which I want to draw attention. One is the special difficulty of measuring the wealth of the wealthy. The other concerns the complexities of defining income and the importance of taking noncash income into account. A category of noncash income deserving particular attention is the implicit income derived from owning and occupying one's own home.

Measuring the Wealth of the Wealthy

The Lamas and McNeil paper presents evidence of excellent response rates being obtained in the first few rounds of SIPP data collection. It also illustrates the first of the general points I want to make, the one about measuring the wealth of the wealthy.

SIPP will provide better data on the asset of lower and middle income groups than on those held by the wealthy. This is not a criticism of SIPP -- a single survey cannot do absolutely everything. Rather, it is a caution to be borne in mind about the need for careful studies to evaluate the quality of the asset data. And it may even be an argument for the occasional use of enhanced samples of upper income households.

There are several reasons for anticipating diminishing accuracy of the asset data as one moves up the income scale. The assets of the wealthy are more varied and held in more complicated forms. The chances of forgetting some assets increase and so do the incentives for concealment. The wealthy are habituated to not telling the government any more than they must.

In addition, the variety of forms in which the assets can be held and the rapidity with which new investment vehicles evolve make it impossible for the SIPP questionnaire to specifically cover them all. The SIPP questions are detailed and comprehensive, but there will always be some new forms of commodity futures, stock options, or other esoteric investments which won't get picked up, or which won't be accurately valued.

The pattern of response rates described in the Lamas-McNeil paper reinforces one's prior expectation of poorer response rates among the wealthy. Overall, response rates promise to be very good. I say "promise", because the Lamas-McNeil data on response rates is based in part on a pre-test, together with an assumption that, if callbacks had been completed in the pretest (which they were not), the information would have been supplied in a high percentage of cases.

Let us suppose, however, that actual response rates equal the Lamas-McNeil projections. Nonresponse among those ages 45 to 64 will be about three times as great as for those under age 34. Nonresponse for college graduates will be twice that of those with less than a high-school education. In other words, the age and education groups known to have the highest average incomes have the highest nonresponse rates.

These response rates pertain to asset ownership. When the questions about asset amounts were asked, nonresponse was especially high regarding stocks and bonds, types of assets held predominately by upper income households.

To sum up, the percentage of total wealth missed by the SIPP survey is likely to be substantially greater than the percentage of nonrespondents to the survey. There will be a need for studies which can quantify this gap.

A further problem with respect to valuing and interpreting assets is that asset values fluctuate. ValueS of some assets, such as common stocks, can fluctuate suddenly and by large amounts. The Dow Jones Industrial Average increased 52 percent from June 1982 to June 1983. This, of course, was a genuine increase in wealth for stock holders and one of which they generally were aware. But the 8 percent rise in the Dow between July 30 and August 2, 1984 may or may not turn out to have any meaningful degree of permanence. Individual stocks can fluctuate by far more than the market averages. How accurately will survey respondents be able (and willing) to value assets such as common stocks? And how meaningful would an accurate valuation as of a specific date be anyway?

Changes in asset values represent changes in wealth, but for assets whose value fluctuates widely and freqently around the underlying trend, some smoothing of short run fluctuations will be desirable for many analytic purposes. This question will be of importance both for cross-sectional comparisons of families who hold their assets in different forms and for analysis of changes in wealth over time. The message for the Census Bureau is that the SIPP data should be made available in a way which allows the user the flexibility of using alternative methods of asset valuation.

Owner-Occupied Housing

For the majority of middle-income families, the family home represents their most substantial investment. In comparing the income and wealth of families which do and do not own their own homes, failure to include the implicit income received from living in an owned home produces misleading comparisons.

This can be illustrated by comparing the incomes of two hypothetical retired couples. Both have identical incomes from pensions and social security. The A's live in their own fully-paid-for home, which has a market value of $100,000. The B's have sold their similar home for $100,000 and invested the proceeds in tax-exempt bonds which provide them with income of $10,000 a year.

On a simple comparison of cash income the B's are $10,000 better off. However, they pay $10,000 per year in rent on their Florida apartment, an expense the A's do not have.

In fact, the A's and the B's have equivalent incomes and equivalent wealth, which they have chosen to utilize in different ways, and a comparison based only on money income is quite misleading.

My example, of course, is greatly oversimplified. I have left out the taxes, maintenance expenses, and so forth that the A's incur living in their owned home. Introduction of these and other complexities would be required in any real life statistical presentation, but this would not alter the need to include the value (income) derived from home ownership in any valid comparison of the relative well-being of renters and home owners.

The Bureau of Economic Analysis and the Bureau of Labor Statistics have wrestled with this problem for years, and have found ways of handling it. BEA includes the imputed value of the shelter provided by owner-occupied homes in the personal consumption expenditure component of GNP and the equivalent imputed income in national income. BLS now uses rental equivalence to estimate home ownership costs for the Consumer Price Index. With the new emphasis on noncash income, the Census Bureau also needs to develop ways of including income from home ownership in its income tables.

This issue was brought to mind by both the Radner and the Vaughan papers. Radner isolates that group of elderly who have both low incomes and limited assets and compares them to the universe of all elderly households. This is a very useful cross-classification, and one would like to see how it would look with an expanded income definition which includes noncash income. Looking only at their cash incomes the financial picture for the low income, low net worth group is grim, indeed. The gap between this group and other elderly households would be narrowed by the inclusion in income of food stamps, medicaid, and other means-tested noncash assistance, but it would be widened by including imputed income from home ownership. If the income concept is widened, as it should be, it should be done comprehensively to include the noncash income received from all major sources, not just from government assistance.

The role of noncash income also needs to be considered in interpreting the Vaughan paper. Vaughan asserts that "the aged require less income than the nonaged to sustain a comparable level of living. For example, they are able to maintain equivalent nutritional status at lower levels of food consumption, they often live in fully paid for owner occupied housing, no longer experience the expenses associated with working and generally a smaller portion of their incomes is devoted to taxes."

If this statement is correct at all, it is correct only because the income definition on which it is based is pre-tax money income. This is the concept for which data is most readily available, but it is not the most appropriate concept for making comparisons of income needs (or, as Vaughan terms them, "family equivalence scales").

The better concept would be after-tax, after-transfer money and non-money income, incuding imputed income from owner-occupied housing. I suspect such a comparison would show that it costs the elerly more, not less, to achieve any given standard of living. Many of these costs are met by Medicare and Medicaid, however, and others by the in-kind return realized by continuing to utilize the house, furnishings, and other consumer durables purchased earlier in the life cycle. It is only because a smaller proportion of their living costs are met from current money income that the elderly may be able to achieve a given standard of living with less "income," narrowly defined.

These three papers on SIPP are tantalizing in their promise of the value of the data becoming available. They also serve to remind us of the conceptual and analytic work still to be done if SIPP data is to be given the most useful and meaningful presentations possible.

SURVEY OF INCOME

AND

PROGRAM PARTICIPATION:

SESSION II

This section is comprised of five papers presented in this session which was sponsored by the Section on Social Statistics.

TOWARD A LONGITUDINAL DEFINITION OF HOUSEHOLDS

David Bryon McMillen and Roger A. Herriot, Bureau of the Census

Introduction and Background

Data collection and analysis in the social sciences generally focus on cross-sectional surveys such as the Current Population Survey (CPS). Consequently, most of our concepts and data analysis tools are structured around point estimates of some phenomenon or characteristic. To the extent that we try to develop longitudinal concepts and measures of social phenomena, that is to say viewing events across time rather than at one point in time, we conflict with these cross-sectional structures. It is the goal of this paper to confront that cross-sectional/longitudinal conflict and attempt some reconciliation. More specifically, this paper attempts to develop longitudinal definitions of households and families which are useful for observing these units across time and for constructing aggregate characteristics across that time period, while not creating serious conflict with our cross-sectional constructs of household type and composition. We begin this exercise by examining cross-sectional household concepts from the CPS and recounting the deficiencies of that perspective. Next we will examine several types of longitudinal definitions, identify the type that is usually cited as most useful, and describe three definitions within that framework. In the third section of this paper we will evaluate the definition options available in terms of utility as well as what is possible given the data at hand to implement such a definition. Next we will illustrate how this definition might be used in calculating aggregate household characteristics and in tabulations of the number of households, household types, and household characteristics.

Point Estimates and Longitudinal Measures

The household definitions used in the CPS serve as adequate measures for the intended crosssectional purposes. Indeed, few people argue that these definitions create a problem when estimating the number of households by type at the time of the survey. However, when those definitions are used in conjunction with other variables, problems begin to develop.

Discussions on measuring annual household income from the CPS center around the retrospective nature of the measure. Household members as of March are asked to recall their income for the previous calendar year, and the income of all members are aggregated to create a household income. The problem centers on the varying lengths of household membership and the unvarying than to provide familiarity. Static definitions, a person is a member of the household for part of the year and, thus, contributes income for only part of the year, that person's entire annual income is included in the household income. Similarly, persons who are not members of the household at the time of the survey are not included in calculating the annual household income, even though they may have contributed income for most of the year. This type of criticism is often used to question the adequacy of the CPS income measures; however, it is better viewed as an example of the problems created by combining a point estimate of household composition with a longitudinal (annual) measure of income. Inevitably, the compromises necessary to combine such cross-sectional and longitudinal constructs produce a less than ideal measure.

Similar criticism of the CPS household data can be made. If we examine consecutive March measures of the distribution of households by type we observe little change. The CPS measure of households masks most of the interesting change in the distribution of house holds. For example, in recent years the number of married couple households has changed at a rate of less than 1 percent a year, or about 200,000 households. Concurrent with that indistinguishable change are over 3 million marriages and divorces, not to mention changes in household type as a result of the death of one member. The small net change creates the appearance of stability, while masking considerable activity. Again, the problem is not so much the inadequacy of the data, but rather the difficulty of measuring longitudinal events with point estimates.

When criticism is leveled against a particular measure, the problem often is not the measure but rather the incongruity between the measuring instrument and the time frame being considered. The examples used above are annual measures, but the same problem exists regardless of the length of time. Most social measurement is discrete while time is continuous. The goal of course is to get to the point where the difference between the two is trivial and can be easily ignored.

In summary, much of the criticism of CPS measures can be attributed to this discrepancy between the time reference of the social measurement and the cross-sectional survey instrument. One solution to the problem is to decrease that difference by repeatedly measuring the phenomenon in question during the year. Those observations can then be aggregated to produce measures which cover a number of time intervals. It is from this perspective that the design for the Survey of Income and Program Participation (SIPP) has developed.

The design of this survey is to interview the household every 4 months over a 2 1/2-year period, and to collect in those interviews monthly data on household composition, income, labor force participation, and a number of other characteristics. Those monthly data can then be aggregated to larger temporal units such as quarterly or annual measures. However, with the idea of aggregating monthly units comes the problem of defining which units should be aggregated across time and which should not. That is to say, which households are the same over the period, which exist at the beginning of the period but not at the end, and which exist at the end but not at the beginning. Without such a definition, aggregating above the person level is impossible.

Defining households across time is an issue that has been debated for several years without resolve; however, it is necessary that the Census Bureau decide which of the many proposed methods will be used for the publication series from the Survey of Income and Program Participation (SIPP). This paper moves one step nearer to that decision by summarizing the proposals on how longitudinal households should be defined and recommending a system to be used. In addition, this paper will begin to identify what conceptual and processing problems remain unsolved given the definition chosen.

Several proposals have been offered for defining longitudinal households. Griffith (1978) outlines six measures, one of which is the traditional Current Population Survey (CPS) definition. Griffith also proposes that the Census Bureau use several definitions in tabulating households from SIPP. Others who had proposed variant measures include two from Davey (1980), Crosby (1979), Lane (1978), & Smith-Ycas (1981). Ycas (1981) in summarizing past work identifies four keys for labeling definitional methods: static; dynamic; staticdynamic hybrids; and attribute methods. In the following section we will discuss several types of longitudinal definitions for households.

Types of Longitudinal Household Definitions

Static definitions of households fix the household composition and characteristics at a given point in time and calculate other attributes from that point. These definitions are the standard cross-sectional perspective on households common to the CPS and other similar surveys. Using a point estimate of household composition, other attributes are calculated assuming that the composition chosen existed for the full period. Thus, some estimate of annual income for each member is aggregated to produce an annual household income, regardless of whether members were there for the full period or joined the day before the interview. This type of household definition is the logical outgrowth of cross-sectional surveys where interviews are conducted at one point in time and aggregates of past events are a function of respondent recall. This type of definition coincides with the instantaneous conception of a household which we use from day to day.

Static definitions are both useful and familiar for cross-sectional surveys; however, they serve little purpose in longitudinal surveys other than to provide familiarity. Static definitions, for a number of reasons, ignore the dynamic activity common to households--households are formed by marriages and dissolved by divorces, children leave home and set up their own household, or move in with relatives, and so on. It is difficult to justify the expense and complexities necessary to measure these dynamics if we then suggest to ignore them in defining households. It is useful to portray static definitions here, however, for they represent one end of the definition spectrum.

Dynamic definitions of households occupy the other end of the spectrum. These definitions recognize change as inherent in observsing households across time, and attempt to incorporate that change into the definition. Thus, household characteristics and attributes become vaiable to measured as households change, are created, and dissolve during the period of observation. In other words, these definitions attempt to minimize the extent to which dynamic concepts are forced into static categories. Needless to say, dynamic definitions are better suited to a longitudinal survey such as SIPP; however, such definitions are difficult to devise and to carry out.

One of the first difficulties encountered with dynamic definitions is that they produce measures which are not readily familiar to many of those who use census data. The most common illustration of this point uses household size. Static definitions produce measures of household size such as 2 or 3, which are inti=uitively meaningful. That is, they fit with our instantaneous image of households because they represent the household size at one point in time--the survey date. Dynamic definitions produce measures of household size which look more like averages across a number of households--2.4 or 3.2 members in the household. These measures are summary statistics of the household experience, summarizing across time. In other words, dynamic definitions force us out of that instantaneous view of households and into thinking about them as something which change across time; our statistics produce a summary of that change. Ycas (1981) suggests several ways of handling the problem of household size-rounding, using modal size, etc.--however, it may be best to reeducate the reader to think of annual household characteristics as the aggregate of a number of discrete experiences. There are other more troublesome problems to be dealt with in developing dynamic household definitions. I will deal here only with definitional problems, acknowledging that there are also measurement problems to be considered later in this paper.

Unlike many demographic variables, there are several aspects of dynamic households for which there is no definition or consensus as to what constitutes a change in type. For example, if a husband and wife divorce, there are several ways we can account for this on our household ledger. We could count this as the dissolution of the husband/wife household and the formation of two new households. This results in a net increase of one household in existence at that point in time and an increase of two households when counting the number of households existing during the period. Alternatively, we could allow one household to be the continuation of the husband/wife household. Again we have a net increase of one household in existence, but because of the continuing household, we increase only by one the number of households existing during the period. To generalize, a household may experience a number of changes across time and we can converse easily about the discrete events. However, we do not have a clear concept of when those changes result in the formation of a new household and the dissolution of an old household. One extreme is to say that any change to the household composition results in a new household. At the other extreme are those who say that this is an issue without resolution and suggest that we abandon the measurement of household characteristics except as they pertain to individuals. In other words, before we can implement a dynamic definition of households, we must first develop a set of continuity rules or accounting principles which identify cases of household dissolution, household formation, and cases where two households at two points in time are identified as the same household.

Most longitudinal household definitions that have been proposed fall somewhere between the static and dynamic extremes. Each acknowledges the difficulty of developing continuity rules, and proposes some static-dynamic blend to finess those problems. A number of cross-sectional/dynamic hybrid definitions have been proposed. One set of these definitions is quasi-dynamic and acknowledges that a set of continuity rules has yet to be developed. Another set is basically a static system designed to avoid the continuity dilemma. Neither of these alternatives is particularly attractive. In the latter case, most of the alternatives create as many problems as they solve. In the former case, if we are going to develop a set of continuity rules, then there is little need for a hybrid definition.

Attribute-type definitions are drawn from the work done on the Panel Study of Income Dynamics (PSID). The goal in these definitions is somewhat different than in the previous discussion. Rather than attempting a longitudinal definition of households, this system calculates a series of cross-sectional households at some smaller time interval, and then ascribes the characteristics of the household to each individual. Measures for some longer time interval are then calculated by aggregrating the series of point estimates across each individual to represent that person's household experience during the period. This system will yield the number of persons who lived in "households" with a monthly income of $1000 to $1500 during the year; however, it is more difficult to derive the number of households with an annual household income of $12,000 to $18,000. In fact, without an additional set of assumptions, this system does not produce an accounting of households across time. In order to develop household statistics within the attribute system it is necessary to assume, for example, that the householder at the end of the period will represent the household experience. Then the household attributes ascribed to that person are aggregated to produce household characteristics. Those aggregrated attributes represent the householder's experience during the period, but not necessarily the experience of the other persons in the household at that time. As can be seen, an assumption such as this contains many of the weaknesses of using a static or cross-sectional definition of households, with few obvious advantages at the household level.

In summary, there are four types of household definitions which have been proposed for use with longitudinal surveys: cross-sectional; dynamic; cross-sectional/dynamic hybrids; and the attribute system. The cross-sectional approach is clearly inappropriate since it ignores the dynamic nature of the data. The attribute system incorporates the dynamic aspects of the data but dodges the issue of developing continuity rules for households. Consequently, this system, by ignoring the social structure of households, produces many of the same problems raised in criticism of the CPS measure of annual household income. It is clear that a dynamic definition is the most desirable alternative, but agreement on just how that definition ought to be formed is elusive.

In the following section, I will discuss dynamic definitions of households and present three sets of continuity rules. The first, proposed by Norton (1982), defines change, rather than continuity, as movement between major types of households. The second is a reciprocal majority rule system developed by Dicker and Casady (1982). The third, developed by Siegel (1982), sets out a set of demographic principles to which continuity rules should conform and then develops a continuity rule within that framework.

A Dynamic Definition of Households

A dynamic definition of households is much easier to describe than to execute. Dynamic household definitions do little more than acknowledge change in household composition or type and determine that the change must be incorporated. In other words, dynamic definitions acknowledge a set of accounts and to some extent set up the framework for those accounts, but usually do not explicate the principles by which the ledger is filled.

For SIPP, it is suggested that we tabulate the changes that occur in household composition and type during the period covered by a panel. That is to say, we must acknowledge that households change, are formed, dissolve, and sometimes stay the same. In the simplest form then, our accounts will record the formation and dissolution of households from our original sample as well as changes in size and type. These tabulations of change will use as often as possible the standard descriptors, such as family and nonfamily households, and the categories associated with those types. Our dynamic definition begins with the static CPS definition defined at the beginning of the panel and then traces the changes that occur to those households across the duration of the panel. Norton (1982) suggests that we define a household as changed when its membership changes in such a manner as to classify it as a different type of household. He proposes that we acknowledge change between the following types of households: 1) married-couple household; 2) male family-household; 3) female family-household; 4) male nonfamily-household; and 5) female nonfamily-household. Thus any change which results in a household falling into a different category results in the dissolution of one household and the formation of another. To illustrate this system consider a husband-wife household which experiences a divorce. Norton's scheme would consider the husband/wife household dissolved and two new households formed. The two new households would be family households if there were children present and nonfamily households otherwise.

A second longitudinal definition has been proposed by Dicker and Casady (1982) for use with the National Medical Care Utilization and Expenditure Survey (NMCUES). In a slight departure from the definitions discussed here, Dicker and Casady focus on families rather than households; however, that does not pose any serious problems. Like others who have approached this problem, Dicker and Casady begin with the realization that there is not consensus on when families begin or cease to exist; rather, such transitions are in part a function of the problem being investigated.

The NMCUES model for defining longitudinal families requires that antecedents and descendent families or, in their terms, predecessor and successor families be defined reciprocally. That is to say, any rules defining relationships across time must be applied to both families simultaneously. Dicker and Casady next demonstrate that when applying these rules you wind up with links between a number of households. That is to say, any family is likely to have more than one predecessor and more than one successor family. Thus, the problem lies in defining which of the possible pairs will be defined as the longitudinal family. As with most longitudinal definitions, the system eventually reduces the decision to what will be defined as the same and what will be defined as change.

Dicker and Casady chose to define sameness by a majority rule. The successor family which receives the majority of members from the predecessor family is identified as the "principal predecessor." These two families then form the linked or longitudinal family unit. Finally, in cases where families split evenly, it should be noted that the NMCUES model does not define a longitudinal unit, but rather dissolves the predecessor family and considers all successor families as newly formed.

Five rules of relationships focus Siegel's (1982) development of household demography. The first two state that a household can have only one descendent and one antecedent that is identified as the same household. That is to say, when a household splits, only one of the subsequent households can be identified as "the same" household. The third and fourth rules identify households which are not the same as some preceding or succeeding household. Households which are not the same as any antecedent household are newly formed; a household not the same as any descendent household is dissolved. The final rule, one of transitivity, states that if A is the same as B and B is the same as C, then A must be the same as C. All that remains to complete this set of accounting principles is a definition of sameness. The rule is offered that two households separate in time and having the same householder are the same household.

Continuity based on following the householder has been criticized because of the somewhat arbitrary way in which the householder is defined, and because it creates what some consider unreasonable change within a continuing household. The most frequently cited example of such change is following the male after a divorce when the children remain with the female. An alternative to Siegel's householder rule which is consistent with his demography of households is to follow the principal person. The principal person is the female in the married-couple household and the householder in all other households: this is the concept used in developing household weights in CPS. By following the principal person, we alleviate the problem cited above. Of course the problem now occurs when the children stay with the male following a divorce, a much less frequent event.

FIGURE 1

	Norton	Siegel	Dicker
Time 1	ABcd*	ABcd*	ABcd*
Time 2	A* Bcd*	A Bcd*	A* Bcd
Time 3	Ac* Bd	Ac Bd	Ac Bd
Time 4	Acd B*	Acd B	Acd B

*New household

Let us consider briefly the strengths and weaknesses of these three systems focusing on two issues: 1) the number of households created across time, and 2) the extent to which the definition promotes or discourages longitudinal analysis. Norton's system comes the closest to maximizing change and, as a result, creates more households than the others. Consider the divorce example from above, but two children remain with the female. Both Siegel and Dicker and Casady would produce a total of two households resulting from the divorce. Norton's system produces three households: 1) the original married-couple household; 2) a male nonfamily-household; and 3) a female family-household. Let us continue following these people and assume that the children leave the female one at a time and join the male (see figure 1). In Norton's scheme, the first move by a child would produce the dissolution of the male nonfamilyhousehold and the creation of a male family-household. Our longitudinal count of households now stands at four. Neither Dicker and Casady nor Siegel would produce new households as a result of the children moving. moving. When the second child moves, the female family-household is dissolved and a female nonfamily-household is created. The male family-household remains unchanged. Over these four observations, Norton's system produces five households; both Siegel and Dicker and Casady produce only two by allowing the continuation of a household across these changes. Let us then look at those continuing households. The continuous household for Siegel's householder rule starts as a four-member married-couple household, dwindles to one member--the male, and increases to two with the addition of one child and then to three with the addition of the second child. On the other hand, Dicker and Casady's continuous household begins as the four-person marriedcouple household and is transformed by the divorce to a three-person female-headed household, then to a two-person and, subsequently, a one-person nonfamily-household. It should be noted that these two continuous households follow opposite courses after the divorce. The continuous household under the principal-person rule would be identical to Dicker and Casady's continuous household.

We should stop at this point and examine yet realistic example of household change. First, if we are interested in counting households (the number or percent of households in poverty during the year, for example), then a continuity system such as Norton's, which allows for continuity in only the most trivial cases, creates a much larger number of households. Suppose the female half of our mythical household was in poverty after the divorce. By Norton's count, during that year we would have 20 percent of the households in poverty. Dicker and Casady and Siegel would show 50 percent of their households in poverty. A second observation to be made here is that, regardless of what sort of continuity rule we adopt, we will observe households which contain a wide variety of change. The question we must ask is whether we accept that households undergo such change and remain intact.

As noted above, each of the three systems has its constituency and its detractors. Siegel's system is criticized because of the disjunctures that can occur following a divorce. For example, the continuous household will follow a male householder who divorces his wife even though the wife and children remain in the housing unit as a group. Similarly, Norton's scheme is criticized because of the lack of attention paid to continuity. Dicker and Casady are criticized by the mechanical nature of the majority rule. Why, it is asked, should one person make all the difference in whether a family is designated new or continuous?

None of the definitions of longitudinal households offered in the literature has proved viable. However, in the process of discussing this issue with several demographers and economists, it became clear that any definition which labels a transition from a family household to a nonfamily household as continuous causes problems; although there are some cases where the transition from nonfamily to family household occurs within a continuous unit. The most obvious of these is the marriage of two persons who have been living together. Drawing on that experience, we determined that we should develop a longitudinal definition of families separate from that for nonfamily households.

We begin with the CPS definition of a family as two or more persons, one of whom is the householder, related by birth, marriage, or adoption, and residing together. To make this cross-sectional definition dynamic, we must add the time dimension or develop a continuity rule. Thus, a longitudinal family is defined as two or more related persons, at least one of whom is the householder or spouse of the householder, who had the same household experience over two or more consecutive months. We further stipulate that no more than one core family unit with children can continue from a previous-month family. Three levels of criteria are offered to distinguish cases where both parents and children split into two or more households. The first-level criterion, for continuity, is that the family with the most child-months is identified as continuous. The second level, to distinguish between families with the same number of child-months, is the family with the most family-months. In the third level, if two potential continuing units tie on both of the above criteria, then the continuing unit will be assigned randomly. Two elements have been added to the CPS definition: 1) the time dimension, and 2) the inclusion of spouse as part of the continuity criteria.

Let us examine this definition more carefully. Consider again our four-step example of divorce and then the movement of two children one at a time from one parent to another. Following the separation, the Bcd family would be the continuing family because it contains two or more members of the initial family, one of whom is the householder or spouse. The A household would be new because of the transition from family to nonfamily status. Following the movement of the first child, c, the Ac family is considered newly formed because of the transition from nonfamily to family status. Finally, the movement of d from the Bd family to the Ac family results in a new nonfamily household, B. Using the notation from figure 1, we have:

1.	ABcd*		3.	Ac*	Bd
2.	A*	Bcd	4.	Acd	B*

Next, we must confront defining continuity for nonfamily households. A nonfamily household is a householder living with nonrelatives only. For these cases, we have adopted a 50-percent rule. As long as the householder and 50 percent or more of the household is the same at two points in time, the household is considered a continuous household. The distinction between this and the majority rule is that, rather than creating new households for even splits, this rule provides for continuity. This definition of family and nonfamily households provides the basic parameters. What remains is to develop a set of programming specifications to implement this definition.

Other possible longitudinal units or groups exist in relation to federally funded support programs. For example, food-stamp units and AFDC units are defined independently of the household and, in fact, households may contain more than one of these units. Longitudinal units for these programs will be defined on the basis of the person in whose name the program application is filed. For example, in a husband-wife, two-child family, the male is defined as the food-stamp recipient. If he leaves, that food-stamp unit is dissolved; a new one is formed if the female reapplies and is found eligible.

Perspectives on Household Characteristics

In this section, I will address the uses of this longitudinal definition of households and argue that we need to tabulate household data from SIPP using at least two types of longitudinal definitions. The need for two types of definitions is a function of the kind of household information needed.

I am arguing that we need to use both a dynamic longitudinal and an attribute-type household definition, because we are interested in both the experience of households and of individuals in households. There are some characteristics like annual household income which we need to examine both as a characteristic of the household and of the individual. This is only to say that there are multiple meanings associated with the concept of income. In CPS, where we have only one way of obtaining income data, we attach all of those meanings to that single measure. SIPP allows us to decompose that measure and look at the components more carefully

To summarize, I have argued that to fully appreciate the household dynamics we observe in SIPP and to portray that activity over a year, we should provide two types of tabulations. The first tabulates household characteristics using a longitudinal definition and examines how changes in some characteristics result in changes in others. The second type of tabulation examines how household characteristics affect individuals across time.

REFERENCES

J.E. Griffith, Impact of a Longitudinal Survey on the Data Base. (ISDP Working Paper #1, 1978.)

E. Davey, Memorandum for the record. 12/24/80.

B. Crosby, Memorandum to Lane, Ycas, and Mahoney. 7/79.

J. P. Lane, Four Alternative Ways of Defining "Annual" Analysis Units. (ISDP Working Paper #6, 1978.)

Smith-Ycas, Developed by Ycas based on reports of proposal by J. Smith. 1981.

M. A. Ycas, Income Analysis Units for Longitudinal Files. 1981.

A. J. Norton, Notes prepared for 11/18/82 meeting of SIPP Household Task Force.

P. Siegel, Notes prepared for the 10/81 meeting of the Census Advisory Committee on Population Statistics.

M. Dicker & R. J. Casady, A Reciprocal Rule Model for A Longitudinal Family Unit. Paper presented at the 1982 meetings of the American Statistical Association.

LIFETIME WORK EXPERIENCE AND ITS EFFECTS ON EARNINGS: DATA FROM THE INCOME SURVEY DEVELOPMENT PROGRAM

John M. McNeil, U.S. Bureau of the Census and Joseph J. Salvo, NY Department of City Planning

The extent to which persons remain attached to the labor force over the course of their working-age years has important economic and social implications. Differences in labor force attachment between men and women has been cited as one major reason why women earn less than men. This study presents data from the 1979 Income Survey Development Program (ISDP) on lifetime work interruptions and examines the relationship between work interruptions and earnings. Descriptive data showing the extent to which men and women have experienced work interruptions are presented, followed by an analysis of the impact of work interruptions on earnings. The study concludes that work interruptions explain only a small proportion of the earnings differential between men and women.

The 1979 ISDP was a panel survey of approximately 9,000 households that were visited at 3-month intervals over a period of a year and a half beginning in February 1979. The survey, part of the development stage of the new income survey called the Survey of Income and Program Participation (SIPP), was a joint effort of the U.S. Department of Health and Human Services and the U.S. Bureau of the Census. The third wave questionnaire contained a section on personal history and within that section were questions on lifetime work interruptions. The questions (reproduced in Figure A) asked whether the person had ever been away from work for 6 months or longer for each of three reasons: (1) because he or she was not able to find work, (2) because he or she was taking care of home or family, and (3) because he or she was ill or disabled. Beginning and ending dates were recorded for each interruption. A maximum of four interruption periods were identified for each of the three possible reasons for interrupting.

A major reason for the interest in data on lifetime work experience is the desire to use such data in the analysis of male-female earnings differentials. The tenets of human capital research have traditionally stressed the importance of work experience patterns as a determinant of earnings. The descriptive data presented in the first part of this report confirm that the lifetime labor force attachment of women is weaker than that of men. Because of interruptions for familial reasons, women have a much higher overall rate of work interruptions than men and they spend a much higher proportion of their potential work years out of the labor force. Such findings have led at least some social scientists to posit that traditional familial responsibilities are one major reason why women earn less than men. This section will describe selected studies of the relationship between work interruptions and earnings and will present an analysis based on the 1979 ISDP data.

Previous Research

A major constraint in early efforts to examine the relationship between experience and earnings was the lack of data on lifetime work experience. More recently, however, a number of studies have been published which exploit the important data which has been made available from the National Longitudinal Surveys of Labor Market Experience (NLS) and the Michigan Panel Survey of Income Dynamics (PSID).

Suter and Miller (1973) were among the first to analyze the retrospective work history data from the NLS. They studied a cohort of women who were 30 to 44 years of age in 1967. Work experience was based on a question which asked about the total number of years in which the person 3 had worked at least 6 months. Suter and Miller concluded that there was a close association between earnings and length of work experience among this cohort of women.

Mincer and Polachek (1974) extended the analysis of the NLS retrospective data. They specified two reasons why discontinuous work history patterns might lead to lower earnings. First, interruptions in market work lead to lower levels of accummulated human capital. Second, interruptions cause a depreciation in existing human capital. That is, time spent away from market work has a cost beyond the effect of foregone experience. In their analysis, Mincer and Polachek found that the amount of time spent at home had a negative impact on earnings even when experience was also included in the earnings equation. They concluded from their analysis that a depreciation effect does, in fact, exist.

This finding was challenged by Sandell and Shapiro (1978) on the grounds that the NLS data used by Mincer and Polachek were subject to various coding errors. They replicated certain of the Mincer-Polachek research using a corrected NLS file and concluded that the original study had overestimated the depreciation effect.

Corcoran (1979) conducted an analysis of the effect of experience and interruptions on earnings using retrospective data from the PSID. One of the major advantages of the PSID data set was that the sample, in contrast to the NLS samples, was representative of the female population 18 to 64 years of age. Corcoran found very little evidence of a depreciation effect. There was no effect for White women and only a minor effect for Black women. Corcoran also argued that restricting the analysis group to women 30 to 44 years of age is likely to overestimate depreciation because many women in this group have recently reentered the labor market and are likely to be affected by misinformation about job opportunities.

More recently, Mincer and Ofek (1982) used NLS data for 30- to 44-year-old married women to reaffirm the depreciation hypothesis. In an analysis of longitudinal (rather than retrospective) data from the NLS, they found that reentry wage rates were lower than wage rates at the time of labor force withdrawal. Furthermore, longer interruptions carried greater wage penalties. They also found, however, that wage rates tended to grow rapidly upon return to work. The observed amount of depreciation, they concluded, is dependent upon the length of the interruption and the length of time spent back in the labor force.

ISDP Data

The effect of work interruptions on earnings was examined by using the data described earlier to construct variables representing interruptions and experience. These variables were included in regressions which related hourly earnings to a set of explanatory variables. The universe for this part of the study consisted of all persons 21 to 64 years of age with wage and salary income during the quarter preceding the interview. Separate regressions were run for men and women, with the log of hourly earnings as the dependent variable.[1]

The interruption and experience variables used in the regressions include the following:

UNEMP = 1 if person had ever experienced an interruption due to an inability to find a job; 0 otherwise.
DISAB = 1 if person had ever experienced an interruption due to illness or disability; 0 otherwise.
TIME-AWAY = Duration of all interruptions[2] as proportion of potential work years.[3]
EXPER = Number of potential work years minus minus duration of all interruptions.[4]
EXPERSQ = The square of EXPER
FT = 1 if the jobs the person has worked at have usually or always been full-time jobs; 0 otherwise.

The interruption variables were specified in the above form because it was hypothesized that earnings could be affected by the existence of an interruption as well as by the length of an interruption. Because interruptions due to unemployment or disability had a relatively small effect on the proportion of potential work years spent away from work, they were entered as zero-one dummy variables. Because interruptions for familial reasons had a very strong effect on the amount of time spent away from work, they were allowed to enter the equation through their effect on the TIME-AWAY variable. The general experience variable, EXPER, was entered in its own form as well as in its squared form, EXPERSQ. The inclusion of the squared form was intended to capture the nonlinear effect of experience on earnings. (The returns to experience tend to flatten after some point.)

The education variables included in the regression were designed to take advantage of the ISDP personal history questions on highest degree obtained, vocational training, and types of courses taken in high school. They included the following:

EDUC1 = With an advanced degree
EDUC2 = With a bachelors' degree
EDUC3 = High school graduate (reference group)
EDUC4 = Not a high school graduate, with a vocational training certificate
EDUC5 = Not a high school graduate, no vocational training certificate
COURSES = Number of selected academic courses completed in high school

Finally, a set of variables representing marital history were included:

MARR1 = Married, no marital disruption (reference group)
MARR2 = With a marital disruption (ever widowed, divorced or separated)
MARR3 = Never married

The basic results of the survey are presented in tables A and B, the means for all variables are shown in table C and the regression results are shown in table D. Results are shown for White women and men as well as for all women and men in order to facilitate comparisons with previous studies. Results are also shown for men and women 30 years of age and over with no familial interruptions as an alternative method of examining the influence of work interruptions.

The large differences between the sexes in the degree of work attachment are highly visible in table C. Men had, on the average, about 19 years of work experience and had spent only about 2 percent of their potential work years away from work. Women, on the other hand, had 14 years of work experience and had spent about 20 percent of their potential work years away from work. There were small or insignificant differences between men and women in the mean values of the other experience and interruption variables and in the mean values of most of the education and marital history variables.

Men, however, were more likely than women to have received advanced degrees and a larger proportion of women than men experienced. The average hour earnings of all women was $4.38, about 63 percent as high as the average hourly earnings of $6.92 for all men.

The regression results confirm the importance of experience as a determinant of earnings. The general experience variables EXPER and EXPERSQ are highly significant for both men and women and are important relative to other variables in the determination of hourly earnings. Attachment to fulltime work also has a significant effect on earnings. The coefficients of the experience variables show that the returns to experience are greater for men than for women.

The interruption variables, in general, have a negative effect on earnings, but the effect is not particularly strong or consistent. The coefficient of TIME-AWAY is significant for both men and women in the equation for persons of all races, but is significant for women only in the equations for White men and women. Interruptions due to illness or disability (DISAB) have a significant negative effect on earnings in five of the equations, but interruptions due to inability to find work have a significant negative effect in only two of the equations.

That an earnings equation contains both experience and interruption variables that are significant is evidence that a depreciation effect does exist. In the equation for men of all races, the experience variables EXPER, EXPERSQ, and FT are highly significant and the interruption variables UNEMP and DISAB are also significant. In the equations for women of all races, the experience variables and the interruption variable TIME-AWAY have highly significant effects on earnings. The conclusion is that a depreciation effect does exist and information about work interruptions will improve those models which attempt to explain earnings.

The coefficents of the education and marital history variables are in line with expectations, but two findings should be noted. First, the coefficient of EDUC4 for men is less negative than the coefficient of EDUC5. This finding suggests that a vocational training certificate has a positive effect on earnings. Second, the coefficient of COURSES is highly significant even though other measures of educational attainment are also present in the equation. That is, for the purpose of explaining earnings, it is important to know about the types of courses taken in high school even when we already have information about years of school completed and highest degree obtained.

Table C shows that the mean earnings of women are only about 62 percent of the earnings of men even when the group under study is comprised of persons 30 years of age and over with no familial interruptions. This differential exists even though women in this universe have approximately the same mean years of experience as men. Table D shows why the large differential exists even when the mean values of experience are so close. Among the men in this group, the coefficient of EXPER is highly significant, but among women, the coefficient is not significant.

In general, standardized regression coefficients reveal that the work interruption variables are less important than either the general experience or education variables as determinants of earnings. This holds true for both men and women. So, while the work interruption variables do show that a depreciation effect exists, general work experience and education are more critical determinants of earnings.

The earnings equations which have been developed for this report can be used to examine the extent to which differences in work history (experience and interruptions) are related to the earnings gap between men and women. That is, given the coefficients of their own equation, what would the earnings of women be if they had the same mean values as men for the variables measuring experience and interruptions. Table E shows that the effect of assigning to women the mean experience and interruption values of men is to reduce the earnings gap by only 12 percent.

Problems in retrospective measures of work experience and work interruptions

One of the goals of SIPP is to develop a data base that can be used to investigate the relationships among income, program participation, and personal history including work history. A certain amount of work history will be obtained as persons are followed during their time in the panel, but persons spend only 2 1/2 years in the panel. Some work history data may be obtained by matching survey records with Social Security earnings records, but matching records takes time and the amount of work history data that can be obtained from Social Security records are limited. Until 1978, the Social Security record contained information on earnings during a quarter which were subject to the Social Security tax. Therefore, if a person's earnings met the Social Security tax limit in the first quarter of the year, no earnings data would appear for the remaining quarters. Since 1978, the record contains annual data on covered and noncovered earnings.

When the personal history supplement was designed for the third wave of the 1979 ISDP, the problem was to develop a set of questions that could be completed in 2 or 3 minutes and that would provide an indication of lifetime work attachment. The approach adopted was to attempt to identify periods lasting 6 months or longer when the person did not work. The ISDP work history questions are reproduced in Figure A.

There are obviously very great problems in trying to measure lifetime work experience in a brief set of questions. The data from these questions do have a considerable amount of face validity, but it seems reasonable to suppose that the data are also characterized by response problems. One way of identifying possible problem areas is to cross-classify current age by age at first reason-specific interruption. If there is no significant memory loss, then one would expect that the proportion of persons reporting that a first reason-specific interruption took place while they were in a particular age interval would be independent of their current age. Since a cross-classification shows that memory loss is a significant factor in the reporting of first interruptions due to an inability to find work. Persons 21 to 29 were much more likely than older persons to report that such an interruption occurred before their 25th birthday. There is some evidence of memory loss in the reporting of first-time interruptions due to disability, but not to the same degree as interruptions due to an inability to find work. There is no evidence of memory loss in the reporting of first-time interruptions of female interruptions for familial reasons. (The above conclusions are based on the assumption that the age groups had similar experiences.)

The ISDP results were taken into consideration when the time came to design the SIPP questions on work history. In an effort to reduce the problem of memory loss, respondents were asked to begin with the earliest 6 month interruption and work forward. The sequence also attempted to determine the total number of interruptions, then, for each period of interruption, determine the duration of and reason for the interruption. Because the SIPP sequence asks for the total number of interruptions and contains a "Don't Know" box for duration of interruption, we expect to be able to do a better job of imputing for item nonresponse.

Future work

We have finished the field collection operation for the third wave of SIPP, the wave containing the work history data and are in the process of designing processing specifications so that a file can be prepared which contains no item nonresponse.

There are some differences between the work history data collected in ISDP and the SIPP work history data. First, the SIPP sample size is 20,000 households, about twice the size of ISDP. Second, the SIPP data on beginning and ending dates of interruptions should be more complete than similar data from ISDP. Third, unlike the ISDP third wave file, the SIPP file will have data on job and occupation tenure. The SIPP file should be somewhat more useful than the ISDP file, and should allow users to expand the analysis by considering other variables (e.g. job and occupation tenure) and by considering the timing of work interruptions not just their existence and duration.

FOOTNOTES

1 Hourly earnings were calculated by dividing total earnings for the 3-month period by the total number of hours worked.

2 A maximum of four interruption periods could be identified for each of three possible reasons for interrupting.

3 Potential work years were defined as age minus years of school completed minus 6.

4 The ISDP data on employer-specific or job-specific measures of work experience (e.g., tenure with most recent employer/at most recent job) were collected in the fifth wave of the survey and were not available for this study.

REFERENCES

[1] Corcoran, Mary. "Work Experience, Labor Force Withdrawals and Women's Wages: Empirical Results Using the 1976 Panel of Income Dynamics." Women in the Labor Market, pp.216-245. Edited by Cynthia B. Lloyd, Emily S. Andrews, and Curtis L. Gilroy. New York: Columbia University Press, 1979.

[2] Mincer, Jacob, and Ofek, Haim. "Interrupted Work Careers: Depreciation and Restoration of Human Capital." The Journal of Human Resources 17 (Spring 1982): 3-24.

[3] Mincer, Jacob, and Polachek, Solomon. "Family Investments in Human Capital: Earnings of Women." Journal of Political Economy 82 (March/April 1974): S76-108.

[4] Polachek, Solomon William. "Potential Biases in Measuring Male-Female Discrimination." Journal of Human Resources 10 (Spring 1975): 205-229.

[5] Sandell, Steven H., and Shapiro, David. "An Exchange: The Theory of Human Capital and the Earnings of Women: A Reexamination of the Evidence." Journal of Human Resources 13 (Winter 1978): 103-117.

[6] Suter, Larry E., and Miller, Herman P. "Income Differences Between Men and Women." American Journal of Sociology 78 (January 1973): 962-974.

Table A. Work Interruption History by Race, Spanish Origin, and Selected Characteristics: Males

Characteristic	Total (in thousands)	Percent with one or more interruptions lasting 6 months or more due to--				Mean percent of potential work years spent away from work for reasons surveyed	
		All reasons surveyed	Inability to find work	Family reasons	Illness or disability	Value	Standard error
Males 21 to 64 years who ever worked	55,828	25.7	17.3	1.5	10.6	3.3	0.1
RACE AND SPANISH ORIGIN[1]							
White	49,381	24.2	15.2	1.2	10.7	2.9	0.1
Black	5,627	40.2	35.0	3.9	10.7	6.9	0.3
Spanish origin	3,220	34.9	22.7	1.2	15.8	3.3	0.3
YEARS OF SCHOOL COMPLETED							
Less than 12	14,171	40.1	24.9	1.9	20.3	4.7	0.2
12 to 15	29,761	24.7	17.3	1.3	9.3	3.2	0.1
16 and over	11,896	11.0	7.9	1.6	2.4	1.7	0.2
AGE BY YEARS OF SCHOOL COMPLETED							
21 to 29 years	16,048	20.5	18.0	1.4	3.4	4.5	0.2
Less than 12	2,314	40.7	35.5	2.2	7.4	6.8	0.5
12 to 15	10,104	20.8	18.5	1.4	3.2	4.7	0.2
16 and over	3,630	6.9	5.5	.6	1.2	2.4	0.4
30 to 44 years	19,106	23.4	16.2	1.6	8.5	2.6	0.1
Less than 12	3,809	36.7	24.5	2.1	18.2	4.1	0.4
12 to 15	10,278	25.2	17.8	1.5	8.0	2.8	0.2
16 and over	5,019	9.6	6.5	1.4	2.0	.9	0.2
45 to 64 years	20,674	11.9	17.7	1.5	18.2	3.0	0.1
Less than 12	8,049	41.6	22.1	1.7	25.1	4.3	0.2
12 to 15	9,378	28.5	15.6	.8	17.1	2.1	0.2
16 and over	3,247	17.8	12.7	3.0	4.3	2.1	0.3
OCCUPATION GROUP OF USUAL JOB							
Professional, technical, or managerial	15,040	14.7	10.2	1.7	5.3	2.3	0.2
Sales or clerical	6,621	20.6	13.8	.9	7.3	2.3	0.2
Craftsmen	12,825	28.8	18.7	1.4	13.5	2.9	0.2
Operatives	10,254	32.5	20.8	.8	14.9	3.9	0.2
Laborers	5,832	37.9	27.6	2.2	13.6	5.6	0.3
Service	3,457	25.5	14.8	2.3	11.1	4.1	0.4

1/Persons of Spanish origin may be of any race.

Table B. Work Interruption History by Race, Spanish Origin, and Selected Characteristics: Females

Characteristic	Total (in thousands)	Percent with one or more interruptions lasting 6 months or more due to--				Mean percent of potential work years spent away from work for reasons surveyed	
		All reasons surveyed	Inability to find work	Family reasons	Illness or disability	Value	Standard error
Females 21 to 64 years who ever worked	57,258	71.9	14.2	64.1	9.2	30.9	0.2
RACE AND SPANISH ORIGIN[1]							
White	49,812	73.0	12.4	66.8	8.3	32.7	0.2
Black	6,402	63.1	27.4	43.8	17.5	17.6	0.4
Spanish origin	3,014	75.0	23.6	62.4	12.9	27.6	0.6
YEARS OF SCHOOLING							
Less than 12	13,740	79.5	21.7	68.5	20.1	33.5	0.3
12 to 15	34,805	73.3	12.7	66.3	6.6	31.5	0.2
16 and over	8,713	54.3	8.6	48.6	2.6	24.2	0.4
AGE BY YEARS OF SCHOOL SCHOOL COMPLETED							
21 to 29 years	16,804	53.1	17.0	42.5	3.5	20.7	0.3
Less than 12	1,948	70.6	23.2	61.7	5.9	30.9	0.8
12 to 15	11,650	56.8	18.4	44.9	4.1	22.3	0.3
16 and over	3,206	29.1	8.3	22.2	.1	8.9	0.5
30 to 44 years	19,445	77.5	12.3	72.3	6.6	34.3	0.3
Less than 12	4,060	79.8	20.4	73.6	12.3	34.2	0.6
12 to 15	12,366	79.8	9.9	75.3	5.7	34.8	0.3
16 and over	3,018	65.1	11.4	58.5	2.9	32.2	0.7
45 to 64 years	21,011	81.7	13.8	73.8	16.1	35.8	0.3
Less than 12	7,733	81.5	22.0	67.6	27.8	33.7	0.4
12 to 15	10,789	83.7	9.8	79.0	10.3	37.7	0.4
16 and over	2,489	73.8	5.6	70.6	5.4	34.3	0.8
OCCUPATION GROUP OF USUAL JOB							
Professional, technical, or managerial	11,723	61.0	9.5	55.4	5.4	24.4	0.3
Sales or clerical	23,782	75.2	10.7	69.4	5.9	33.8	0.2
Operatives	8,447	74.8	22.2	62.9	14.5	29.4	0.4
Laborers	950	78.3	21.9	67.8	15.1	39.7	1.2
Service	10,543	74.2	19.7	63.4	16.5	32.1	0.4

1/Persons of Spanish origin may be of any race.

Table C. Mean Values of Regression Variables

Variable	Males	Females	White males	White females	Males 30 and over with no familial interruptions	Females 30 and over with no familial interruptions
Interruption and experience variables:						
UNEMP	.135	.117	.121	.098	.137	.117
DISAB	.060	.049	.062	.040	.078	.093
TIME-AWAY	.017	.203	.015	.220	(X)	(X)
EXPER	19.256	14.334	19.337	13.868	25.223	24.644
EXPERSQ	535.421	331.227	537.208	312.682	755.661	745.445
FT	.969	.858	.968	.853	.990	.928
Education variables:						
EDUC1	.069	.044	.071	.045	.080	.076
EDUC2	.161	.128	.166	.135	.142	.107
EDUC4	.019	.015	.017	.009	.023	.020
EDUC5	.204	.157	.191	.145	.243	.274
COURSES	2.170	2.129	2.202	2.167	2.059	1.843
Marital history variables:						
MARR2	.214	.303	.204	.277	.255	.388
MARR3	.154	.172	.144	.159	.057	.203
Earnings variables:						
Hourly earnings	$6.92	$4.38	$7.14	$4.42	$7.58	$4.71
Log of hourly earnings .	1.934	1.478	1.966	1.487	2.026	1.550
Unweighted N	3,157	2,416	2,854	2,101	2,145	593

X Not applicable.

Table D. Coefficients of Regression of Log of Hourly Earnings on Specified Explanatory Variables

(Standard errors in parentheses)

Variable	Employed males	Employed females	Employed White males	Employed White females	Males 30 and over with no familial interruptions	Females 30 and over with no familial interruptions
UNEMP	-.039 (.018)	.002 (.018)	-.028 (.021)	.002 (.021)	-.078 (.018)	.014 (.041)
DISAB	-.125 (.023)	-.040 (.028)	-.144 (.025)	-.088 (.032)	-.143 (.025)	-.183 (.044)
TIME-AWAY	-.312 (.122)	-.128 (.025)	-.068 (.145)	-.155 (.028)	(X) -	(X) -
EXPER	.03515 (.00175)	.02278 (.00184)	.03791 (.00189)	.02495 (.00200)	.03382 (.00306)	.00937 (.00600)
EXPERSQ	-.00058 (.00005)	-.00042 (.00005)	-.00065 (.00005)	-.00046 (.00005)	-.00056 (.00005)	-.00014 (.00012)
FT	.216 (.032)	.112 (.016)	.254 (.035)	.099 (.018)	.363 (.064)	.372 (.048)
EDUC1	.336 (.023)	.358 (.028)	.338 (.023)	.322 (.030)	.327 (.028)	.301 (.053)
EDUC2	.179 (.016)	.218 (.018)	.181 (.018)	.209 (.021)	.231 (.021)	.260 (.046)
EDUC4	-.069 (.039)	-.146 (.048)	-.002 (.044)	-.120 (.067)	-.026 (.044)	-.415 (.092)
EDUC5	-.195 (.016)	-.190 (.018)	-.173 (.016)	-.179 (.018)	-.185 (.018)	-.244 (.035)
COURSES	.038 (.005)	.044 (.005)	.034 (.005)	.052 (.005)	.045 (.005)	.070 (.009)
MARR2	-.023 (.014)	.016 (.014)	-.038 (.014)	.038 (.014)	-.009 (.016)	-.035 (.030)
MARR3	-.192 (.016)	-.009 (.018)	-.141 (.018)	-.008 (.018)	-.279 (.030)	.029 (.035)
Constant	1.318	1.112	1.282	1.098	1.172	.993
R^2	.24	.18	.22	.19	.20	.28

- Represents zero.
X Not applicable.

Table E. Percent of the Male-Female Earnings Gap Explained By Experience, Interruption, and Education Variables

Variables	Employed persons all races	Employed persons White	Employed persons 30 years of age and over with no familial interruptions
Hourly earnings of men	$6.92	$7.14	$7.58
Hourly earnings of women	$4.38	$4.42	$4.71
Earnings gap	$2.54	$2.72	$2.87
Hourly earnings of women adjusted for experience and interruption variables[1]	$4.69	$4.76	$4.85
Difference between actual hourly earnings of women and hourly earnings adjusted for experience and interuption variables	$ .31	$ .34	$ .14
Percent of earnings gap explained by experience and interruption variables	12.2	12.5	4.9
Hourly earnings of women adjusted for experience, interruption, and education variables[2]	4.75	4.80	5.64
Difference between actual hourly earnings of women and hourly earnings adjusted for experience, interruption, and education variables	.37	.38	.93
Percent of earnings gap explained by experience, interruption, and education variables	14.6	14.0	32.4

1 Hourly earnings of women if women had the same mean values as men for experience and interruption variables. Experience and interruption variables include UNEMP, DISAB, TIME-AWAY, EXPER, EXPERSQ, and FT.

2 Hourly earnings of women if women had the same mean values as men for experience, interruption, and education variables. Education variables include EDUC 1, EDUC 2, EDUC 4, EDUC 5, and COURSES.

PANEL SURVEYS AS A SOURCE OF MIGRATION DATA

Donald C. Dahmann, U.S. Bureau of the Census

PREVIOUS USES OF PANEL SURVEYS FOR GEOGRAPHICAL MOBILITY RESEARCH

Panel surveys, the technique of using repeated interviews with a sample of individuals that remains constant, have been employed as a research strategy to assist in understanding processes with an inherent capacity for change for at least half a century. Virtually all of the earliest uses of panel analysis were employed to investigate change in opinions, and particularly in preferences for political candidates (Levenson, 1978; Rice, 1928; Lazarsfeld and Fiske, 1938; and Berelson, Lazarsfeld, and McPhee, 1954). From these beginnings, various forms of panel analysis, and more generally longitudinal analysis, have come to be utilized in a wide variety of social science research applications, and recently were singled out by the Office of Federal Statistical Policy and Standards as an approach to data collection and analysis to which "much attention should be devoted...in the development of [federal] statistical programs for the 1980's" (U.S. Department of Commerce, 1978; p. 321).

The real blossoming of panel surveys, both in terms of numbers of surveys and their size, occurred during the past two decades (and principally during the 1970s). They were employed as a means of monitoring and evaluating the effects of a variety of federally sponsored large-scale experiments and programs, and to enhance our understanding of the factors involved in a variety of social and economic processes, e.g., educational attainment, socialization, and labor force participation.

The principal federally sponsored experiments including panel analysis components were the several Income Maintenance Experiments, the various Experimental Housing Allowance programs, and the Urban Homesteading Demonstration. Uses of panel analysis in each of these programs for geographical mobility research are discussed in the next section.

Several large-scale panel surveys were also initiated during this period in response to the need for basic information on educational attainment, labor force participation, and social stratification processes (Borus, 1982). Major surveys in this second group include the National Longitudinal Survey of Laobr Market Practices, begun in 1966 and sponsored by the U.S. Department of Labor with distribution through Ohio State University's Center for Human Resource Research (popularly referred to as the Parnes data; Parnes, 1974; Center for Human Resources Research, 1974; Bielby, Hawley, and Bills, 1977; Parnes and Rich, 1980; Leigh, 1982; Daymont, 1983); the Panel Study of Income Dynamics, conducted by the University of Michigan's Institute for Social Research for the U.S. Department of Health and Human Services since 1968 (Duncan and Morgan, 1982); Project TALENT, conducted by the American Institutes for Research (Wise and Steel, 1980); Youth in Transition Project, conducted by the University of Michigan's Institute for Social Research (Bachman and O'Malley, 1980); National Longitudinal Survey of High School Seniors, conducted by Research Triangle Institute for the National Center for Education Statistics (Eckland and Alexander, 1980); the Wisconsin Longitudinal Study (Sewell and Hauser, 1980); and Explorations in Equality of Opportunity, initiated by the Educational Testing Service (Alexander and Eckland, 1980).

Questions that these latter surveys were designed to address rarely included geographical mobility as a major topical area. Much attention in terms of questionnaire design or questions asked however, was not required as geographical mobility data flowed as a natural outcome of the follow-up of panel members. Thus, general information on geographical mobility derived of the panel design was sufficient for relating movement with labor force participation, educational attainment, social and occupational stratification processes, household changes with passage through the life course, shifts in housing consumption, etc.

The Experimental Housing Allowance Programs and the Income Maintenance Experiments, on the other hand, were both specifically interested in geographical mobility as an integral element of evaluating the effectiveness of the programs. Both programs, for instance, were concerned with the effect upon, and role of residential mobility in relation to patterns of housing consumption, and the Income Maintenance Experiments were further concerned with the role of migration as a response to income guarantees among lower-income households. The fact that information on the various forms of geographical mobility flowed directly from the panel design of these surveys without specific geographical mobility questionnaire items demonstrates the capacity of panel surveys to assist our understanding the role of geographical mobility in a variety of social and economic circumstances.

Experimental Housing Allowance Program

The Experimental Housing Allowance Program, initiated by the Housing and Community Development Act of 1970, was undertaken to establish empirical evidence of the effects of housing allowances, and of the transfer of small amounts of unrestricted funds to lower-income households on housing consumption. The following questions highlight some of the experiment's major research goals. Would households spend the money on housing? Would the money be used to improve conditions of their current dwelling? Would households move to other neighborhoods? What would be the local housing market impact of such an infusion of funds. To answer such questions, three programs—Housing Allowance Supply Experiment, Housing Allowance Demand Experiment, and Administrative Agency Experiment—eventually enrolled more than 30,000 households at twelve sites throughout the country at a cost in excess of $160 million (Friedman and Weinberg, 1982; 1983; Bradbury and Downs, 1981; Struyk and Bednick, 1981). Evaluation of these individual programs produced several longitudinal analyses, but only one that utilized individual households as its unit of observation. Other longitudinal analyses took housing units and neighborhoods as their units of observation (Hillestad and McDonald, 1983), or obtained retrospective geographical mobility data for individuals (McCarthy, 1983). Individuals in both control and test groups of the Housing Allowance Demand Experiment were traced for three years to observe, among other things, actual patterns of residential mobility and what changes, in terms of housing consumption and residential dispersion (and therefore desegreation), may have resulted from such moves (Rossi, 1981; Hamilton, 1983).

Similar panel analyses were utilized in evaluating the effects of one other important housing experiment initiated at about the same time, the U.S. Department of Housing and Urban Development's Urban Homesteading Demonstration. This effort, begun as a demonstration and later established as a regular program, transferred HUD-owned properties to local control in 23 cities (U.S. Department of Housing and Urban Development, 1977). Its research agenda included inquiry into the effects of residential mobility on patterns of local housing consumption, this time with specific reference to the displacement of low income households from housing they could no longer afford as the result of the HUD owned properties being returned to the market (Schnare, 1979).

Income Maintenance Experiments

Another of the recent massive social experiments that included panel analysis as a research component is the series of income maintenance experiments, the first large scale social experiments to be conducted in the United States, (begun in 1967)—New Jersey Income Maintenance Experiment (Kershaw and Fair, 1976; Watts and Rees, 1977; Pechman and Timpane, 1975), Rural Income Maintenance Experiment (Bawden and Harrar, 1975; Palmer and Pechman, 1978), Gary Income Maintenance Experiment (Journal of Human Resources, 1979), and the Seattle and Denver Income Maintenance Experiment (Journal of Human Resources, 1980). Each of these four programs addressed various aspects of one basic issue: how much would a nationwide guaranteed income cost, and to what extent would families reduce their labor force participation (and therefore earnings) in response to such payments? As with the massive housing allowance experiments, research agendas of the income maintenance experiments did not include geographical mobility as a specific focus. Nonetheless, the panel designs employed in the evaluations, which traced families (households) over a three-year period, produced their own mobility data. Analyses were undertaken of both of the two major forms of geographical mobility: (1) migration, specifically rates of movement from the experimental site to other labor markets and, (2) residential mobility, change of one's dwelling to consume a different bundle of housing services (quality of dwelling, neighborhood services, etc.).
Data to investigate both of these extremely important outcomes of the decision to move were readily derivable from the panel design of the surveys used to evaluate the experiments.

Findings with regard to migration (spatial adjustments in labor force participation) may be summarized as follows: (1) migration out of the experimental site's labor market was significantly increased for married white males and females but not for married black males and females and, (2) outmigration was to locales with generally lower wage rates and with better living environments. Work hours in the new locations were generally less than previously, suggesting either that persons worked fewer hours in their new locations because of their additional income or that their search for a "satisfactory" job in the new locale took some time (Keeley, 1980).

With regard to residential mobility, it was discovered that (1) households moving to a new address generally improved their housing situation (Wooldridge, 1977; Kaluzny, 1979), and (2) the effects of income assistance as a means of enabling renter households to move into a home of their own were mixed (Wooldridge, 1977, Poirier, 1977).

Panel Study of Income Dynamics

A more archetypical panel study, at least in traditional terms, is the Panel Study of Income Dynamics (PSID) conducted by the Institute for Social Research of the University of Michigan for the U.S. Department of Health and Human Services. Initiated during the same period—Great Society Era of the 1960s—as the Income Maintenance Experiments, this panel survey is now in its 17th year of collecting annual information from a representative national sample of about 6,000 families and 15,000 individuals (Morgan and Smith, 1969). The Survey has produced a massive body of data, a massive array of findings (Duncan and Morgan, 1982,; Morgan, 1974; Duncan and Morgan, 1975-1980; and Hill, Hill, and Morgan, 1981), and even outlived its original sponsoring federal agency, the Office of Economic Opportunity.

Panel Study of Income Dynamics data have been utilized to address questions in both of the two major realms of geographical mobility research. It has also been utilized to consider more basic geographic mobility research questions—timing of moves through the life course, relationships between desires or expectations to move and actual movement, and other similar questions that are intrinsic to the geographical mobility processes. The rich set of personal attribute and attitudinal variables in the Michigan panel has enabled residential mobility research to be framed in behavioral terms, whereby households are seen as possessing specific desires and preferences with respect to moving. Structural elements of the participation system—income levels and purchasing power, housing costs, forced relocation, etc.—in this frame of reference then serve to assist or hinder actual patterns of mobility, and therefore preference fulfillment (Roistacher, 1974; 1975; Goodman, 1974; Duncan and Newman, 1975; 1976; Newman and Duncan, 1979; and Newman and Owen, 1982).

Use of these panel data have enabled researchers to examine the characteristics of migrants in various interregional migration systems during the 1970s (Kim, 1980), patterns and consequences of repeat moves (Newman and Ponza, 1981), and to initiate structuring of general models of mobility (Morgan, 1977). Migration as an act resulting in the readjustments of labor markets has also been explored both in terms of causes, such as the effects of unemployment on movement (DaVanzo, 1978), and of consequences for individuals, in terms of income and occupational change (Harris, 1981).

THE SURVEY OF INCOME AND PROGRAM PARTICIPATION AS A SOURCE OF MIGRATION DATA

The Survey of Income and Program Participation (SIPP), a general purpose, large scale (25,000 household), national representative sample survey, has been undertaken by the U.S. Bureau of the Census primarily to provide: (1) improved data on the economic situation of individuals and households and (2) information on federal and state income transfer and social program participation. Individuals are interviewed every four months for the life of a panel. In the case of the first (1984) panel, this will result in a total of 9 waves of interviews for three--quarters of the panel and 8 waves for the remaining quarter. Initial interviews for the 1984 panel were conducted in October 1983 (with a reference period of July-September 1983); the final wave of interviews for the this panel will be conducted in May 1986 (with a reference period of January--April 1986). Current plans for a second (1985) panel call for a somewhat smaller sample size (20,000 households) and eight waves of interviews, which will begin in January 1985.

The earlier review of geographical mobility research topics explored with data from panel surveys presages the types of research we might expect from SIPP. In terms of duration of the panel, SIPP data will be much like those that were derived of the several large scale experiments—both cover a period not exceeding three years. In terms of geographical mobility research therefore, data from both sources may be used to explore change over only a relatively short period of time. Several SIPP characteristics make it a close relative of the Panel Study of Income Dynamics as well: (1) its sample is national (though larger) rather than being limited to selected settlements, and (2) more waves of interviews are being conducted, which will provide better data for establishing joint incidence of movement with other--life course, employment, etc.—events. The geographical mobility research derived of the panel surveys discussed earlier suggests that attention must be given to two specific aspects of such surveys: first, the periodicity of waves and overall duration of the panel; and second, the substantive nature of data collected during each wave. SIPP is unique in the short span of time between waves--four months. This design characteristic makes it particularly valuable as a means of matching residential shifts and migration with other life events such as marriage, divorce, expansion and contraction of household size in general, loss of job, change in job, and the like. No previous national survey has provided such a fine temporal scale for establishing the joint-incidence of geographical movement with important employment and life events.

The fact that each panel collects data for 2 1/2 years presents both advantages and disadvantages. As SIPP panels will not be maintained for years unto decades, as have the several major panel surveys focusing on changes in labor force participation, educational attainment, and social mobility through the life course, analysis of the role, consequences, and duration of effects of geographical mobility through major stages of the life course is not feasible. Nonetheless, 2 1/2 years (plus the fact that a large number of waves will be conducted) is quite sufficient to establish both immediate and some intermediate-term effects associated with geographical mobility on topics of concern such as the spatial restructuring of labor markets. The duration of SIPP panels also provides a reasonable amount of time to relate the expectations of individuals toward mobility to actual patterns of movement.

Once we have accustomed ourselves to the fact that the act of tracking those who move in a panel survey provides migration data, then what must be considered in addressing geographical mobility questions is the basic substance of the questionnaire administered prior to and following the move. In the case of SIPP we are provided with a wealth of relevant migration-related information: labor force participation and employment, industry and occupation, work history, education, health conditions and disability, household composition, and, of course, income. As the same questionnaires are administered at the same times to nonmovers, the opportunity exists for comparing the situations of movers and nonmovers directly.

In consideration of these several points, what should we be looking to SIPP for in terms of geographical mobility research? First, I think that we can expect better data. For decades the Current Population Survey (CPS) has served as our national metric establishing levels of movement among the various components of the nation's settlement system, among regions, and among subpopulations of the nation's peoples. All CPS geographical mobility data are collected retrospectively—sometimes asking respondents to refer to an event that occurred one year ago, sometimes five years ago. These data, like all retrospective data, are subject to biases introduced by the distorting effects of memory loss, dissonance reduction (rationaliazation), and the like. How does SIPP data compare with CPS data? Will SIPP's multiple waves of data collection enable us to specify the overall effects of repeat movers on mobility statistics in a way that cross-sectional data do not? What will differences between the two survey's geographical mobility data tell us? The fact that information on movement (and non-movement) and a wide array of life events will be collected almost as they occur (specific to within four months) is one of SIPP's best features from the perspective of geographical mobility research. Our capacity to specify the relationships between movement and such events as the loss of a job, termination of the receipt of unemployment benefits, marriage, divorce, etc. has never been better.

A set of supplemental migration questions, which should be administered to all individuals for a least one (preferably early) wave of interviewing, should also be considered. First, respondents should be asked a set of mobility preference questions, to relate desires and expectations of movement with patterns of actual mobility. Secondly, retrospective questions on one's general residential history should be asked so that subsequent moves may be related to previous patterns of mobility and locations.

One further aspect of SIPP's design that should not be overlooked when thinking about geographical mobility research is its capacity to provide information on the locales of origin and destination of movers (and nonmovers as well). The ability of SIPP to provide information on conditions in both the labor markets that migrants leave, and those to which they move, is of particular concern when wishing to understand the spatial differentiation of labor markets and to ascertain the causes of subnational (regional) patterns of employment growth and decline. What, for instance, are the relationships between sending and receiving markets in terms of unemployment rates, wage levels, etc.? Are these structural situations consequential in terms of employment? Are different mechanisms operating for blue collar and white collar migrants that such differences articulate? These are some of the questions that should guide attempts to maximize the utilization of SIPP data for geographical mobility research.

CONCLUSION

The Survey of Income and Program Participation enables us to explore new questions concerning both of the major forms of geographical mobility--residential mobility and migration. With particularly good income and public program participation data, good specification of the timing of movement with significant life events, and (potentially) good market characteristics data, SIPP is ideally suited to address a multitude of housing consumption questions. With good information on participation in federal and state sponsored programs, exceptionally good income data, and (potentially) good information on the characteristics of labor markets, SIPP promises to be an incomparable research tool for questions that have heretofore simply gone unasked regarding migration, and particularly as it relates to readjustments of the spatial dimension of labor markets.

We must also be prepared to take advantage of the serendipitous benefits of timing. In this regard, the availability of SIPP data and recent advances in analytical techniques provide us with opportunities that were nonexistant even a decade ago. With regard to analytical techniques I am thinking particularly of those developed during the 1970s for analyzing categorical data (Goodman, 1978; Bishop, Fienberg, and Holland, 1975;), and their specific application to the analysis of change in mobility and panel data (Hauser, 1979; Duncan, 1981; Fienberg, 1980). The richness of SIPP data provide a wonderful opportunity to fully utilize the analytical advances brought about by these techniques to answer a myriad of geographical mobility questions.

The Great Society programs of the 1960s pushed social scientists as never before to ask questions about American society and its economy. In response to these demands, new and better data were collected, new analytical techniques were developed, and new research agendas established. Much was learned from these efforts about the causes, the roles performed by, and effects of geographical mobility on the nation's economic and social structure. SIPP represents a logical outcome of advances in social science data collection that began in the 1960s and an important new opportunity for geographical mobility research. I invite your comments on ways that we at the Census Bureau can enhance this new survey's utility for answering your geographical mobility questions.

REFERENCES

Abt Associates. 1983. Annual Housing Survey National Longitudinal File [Codebook]. Cambridge, Mass.: Abt Associates.

Alexander, Karl L. and Bruce K. Eckland. 1980. "The 'Explorations in Equality of Educational Opportunity' Sample of 1955 High School Sophomores." In Alan C. Kerckhoff. ed. Research in Sociology of Education and Socialization. Vol. 1. Greenwich: JAI Press. Pp. 31-58.

Bachman, Jerald G. and Patrick M. O'Malley. 1980. "The Youth in Transition Series: A Study of Change and Stability in Young Men." In Alan C. Kerckhoff. ed. Research in Sociology of Education and Socialization. Vol. 1. Greenwich: JAI Press. Pp. 127-160.

Berlson, Bernard, Paul F. Lazarsfeld, and William N. McPhee. 1954. Voting: A Study of Opinion Formation in a Presidential Campaign. Chicago: University of Chicago Press.

Bilsborrow, Richard E. and John S. Akin. 1982. "Data Availability versus Data Needs for Analyzing the Determinants and Consequences of Internal Migration." Review of Public Data 10 (December): 261-284.

Bishop, Yvonne M. M., Stephen E. Feinberg, and Paul Holland. 1975. Discrete Multivariate Analysis: Theory and Practice. Cambridge: MIT Press.

Bawden, Lee and William Harrar. eds. 1976. Rural Income Maintenance Experiment: Final Report. Madison: Institute for Research on Poverty, University of Wisconsin-Madison.

Bielby, William T., Clifford B. Hawley, and David Bills. 1977. "Research Uses of the National Longitudinal Surveys." Special Report No. 18. Madison: Institute for Research on Poverty, University of Wisconsin.

Borus, Michael E. 1982. "An Inventory of Longitudinal Data Sets of Interest to Economists." Review of Public Data Use 10 (May): 113-126.

Bradbury, Katerine and anthony Downs. eds. 1981. Do Housing Allowances Work?

Washington, D.C.: Brookings Institution.
Center for Human Resources Research. 1974. National Longitudinal Surveys Handbook. Columbus: Center for Human Resources Research, Ohio State University.
DaVanzo, Julie. 1978. "Does Unemployment Affect Migration?--Evidence from Micro-data." Review of Economics and Statistics 60: 504-514.
Daymont, Thomas and Paul Andrisani. 1983. "The Research Uses of the National Longitudinal Surveys: An Update." Review of Public Data Use 11: 203-310.
Duncan, Greg J. and James N. Morgan. 1975-1980. Five Thousand American Families--Patterns of Economic Progress. Volumes 3 to 8. Ann Arbor: Institute for Social Research, University of Michigan.
Duncan, Greg J. and James N. Morgan. 1980. "The Incidence and Some Consequences of major Life Events." In Greg J. Duncan and James N. Morgan. eds. Five Thousand American Families--Patterns of Economic Progress. Ann Arbor: Institute for Social Research, University of Michigan. Pp. 183-240.
Duncan, Greg J. and James N. Morgan. 1982. "Longitudinal Lessons from the Panel Study of Income Dynamics. Review of Public Data Use 10: 179-184.
Duncan, Greg J. and Sandra J. Newman. 1975. "People as Planners: The Fulfillment of Residential Mobility Expectations." In Greg J. Duncan and James N. Morgan. eds. Five Thousand American Families--Patterns of Economic Progress. Volume 3. Ann Arbor: Institute for Social Research, University of Michigan. Pp. 279-318.
Duncan, Greg J. and Sandra J. Newman. 1976. Expected and Actual Residential Mobility." Journal of the American Institute of Planners 42: 174-186.
Duncan, Otis Dudley. 1981. "Two Faces of Panel Analysis: Parallels With Comparative Cross-Sectional Analysis and Time-Lagged Association." In Sanuel Linhardt. ed. Sociological Methodology 1981. San Francisco: Jossey Bass. Pp. 281-318.
Eckland, Bruce K. and Karl L. Alexander. 1980. "The National Longitudinal Study of the High School Senior class of 1972." In Alan C. Kerckhoff. ed. Research in Sociology of Education and Socialization. Vol. 1. Greenwich: JAI Press. Pp. 189-222.
Ferber, Robert and Werner Z. Hirsch. 1978. "Social Experimentation and Economic Policy: A Survey." Journal of Economic Literature 16 (December 1978): 1379-1414.
Fienberg, Stephen E. 1980. "The Measurement of Crime Victimization: Prospects for Panel Analysis of a Panel Survey." The Statistician 29: 313-350.
Friedman, Joseph and Daniel H. Weinberg. 1982. The Economics of Housing Vouchers. New York: Academic Press.
Friedman, Joseph and Daniel H. Weinberg. eds. 1983. The Great Housing Experiment. Urban Affairs Annual Review, Volume 24. Beverly Hills: Sage.
Goodman, John L. 1974. "Local Residential Mobility and Family Housing adjustments." In James N. Morgan. ed. Five Thousand American Families--Patterns of Economic Progress. Volume 2. Ann Arbor: Institute for Social Research, University of Michigan. Pp. 79-105.
Goodman, Leo. 1978. Analyzing Qualitative-Categorical Data. Cambridge, Mass.: Abt Books.
Hamilton, William L. 1983. "Economic and Racial /Ethnic Concentration." In Joseph Friedman and Daniel H. Weinberg. eds. The Great Housing Experiment. Urban Affairs Annual Review, Vol. 24. Beverly Hills: Sage.
Harris, Richard J. 1981. "Rewards of Migration for Income Change and Income Attainment: 1968-1973." Social Science Quarterly 62: 275-293.
Hauser, Robert M. 1979. "Some Exploratory Methods for Modeling Mobility Tables and Other Cross-Classified Data." In Karl F. Schussler. ed. Sociological Methodology 1980. San Francisco: Jossey Bass. Pp. 413-458.
Hill, Martha S., Daniel H. Hill, and James N. Morgan . 1981. Five Thousand American Families--Patterns of Economic Progress. Volume 9. Ann Arbor: Institute for Social Research, University of Michigan.
Hillstad, Carol E. and James L. McDowell. 1983. "Neighborhood Change." In Joseph Friedman and Daniel H. Weinberg. eds. The Great Housing Experiment. Beverly Hills: Sage.
Journal of Human Resources. 1979. Issue devoted to "The Gary Income Maintenance Experiment." 14: 431-506.
__________. 1980. Issue devoted to "The Seattle and Denver Income Maintenance Experiments." 15: 463-722.
Kaluzny, Richard L. 1979. "Changes in the Consumption of Housing Services: The Gary Experiment." Journal of Human Resources 14: 496-506.
Keeley, Michael C. 1980. "The Effect of a Negative Income Tax on Migration." Journal of Human Resources 15: 695-706.
Kershaw, David and Jerilyn Fair. eds. 1976. The New Jersey Income-Maintenance Experiment. Vol. 1. New York: Academic Press.
Kim, Joachul. 1980. "Characteristics of Migrants Within the Framework of Current Migration Direction in the United States: Some Evidence from Micro-Data Analysis." Policy Sciences 12: 355-378.
Lazarsfeld, Paul F. and Marjorie Fiske. 1938. "The 'Panel' as a New tool for Measuring Opinion." Public Opinion Quarterly 2: 596-612.
Leigh, Duane E. 1982. "The National Longitudinal Surveys: A Selective Survey of Recent Evidence." Review of Public Data Use 10: 185-201.
Levenson, Bernard. 1978. "Panel Studies." In William H. Kruskai and Judith M. Tanur. eds. International Encyclopedia of Statistics. New York: Free Press. Pp. 683-691.
McCarthy, Kevin F. 1983. "Housing Search and Residential Mobility." In Joseph Friedman and Daniel H. Weinberg. eds. The Great Housing Experiment. Beverly Hills: Sage.
Morgan, James N. ed. 1974. Five Thousand American Families--Patterns of Economic

Progress. Vols. 1 and 2. Ann Arbor: Institute for Social Research, University of Michigan.

Morgan, James N. 1977. "Some Preliminary Investigations for a Model of Mobility." In Greg J. Duncan and James N. Morgan. eds. Five Thousand American Families--Patterns of Economic Progress. Volume 5. Ann Arbor: Institute for Social Research, University of Michigan. Pp. 349-368.

Morgan, James N. and James D. Smith. 1969. A Panel Study of Income Dynamics: Study Design, Procedures, and Forms--1968 Interviewing Year (Wave I) Vol. 1. Ann Arbor: Institute of Social Research, University of Michigan.

Nelson, Dawn, David McMillan, and Daniel Kasprzyk. 1984. An Overview of the Survey of Income and Program Participation. SIPP Working Paper Series No. 8401. Washington, D.C.: U.S. Bureau of the Census.

Newman, Sandra J. and Greg J. Duncan. 1979. "Residential Problems, Dissatisfaction, and Mobility." Journal of American Planning Association 45: 154-166.

Newman, Sandra J. and Michael S. Owen. 1982. Residential Displacement in the U.S., 1970-1977. Ann Arbor: Institute for Social Research, University of Michigan.

__________ and Michael Ponza. 1981. "The Characteristics of Housing Demand in the 1970s: A Research Not." In Martha S. Hill, Daniel H. Hill, and James N. Morgan. eds. Five Thousand American Families--Patterns of Economic Progress. Ann Arbor: Institute for Social Research, University of Michigan.

Palmer, John L. and Joseph A. Pechman. eds. 1978. Welfare in Rural Areas: The North Carolina-Iowa Income Washington, D.C.: Brookings Institution.

Parnes, Herbert S. 1974. "The National Longitudinal Surveys: New Vistas for Labor Market Research." American Economic Review 65: 244-249.

__________ and Malcolm C. Rich. 1980. "Perspectives on Educational Attainment from National Longitudinal Surveys of Labor Market Behavior." In Alan C. Kerckhoff. ed. Research in Sociology of Education and Socialization Vol. 1. Greenwich: JAI Press. Pp. 161-188.

Pechman, Joseph A. and P. Michael Timpane. eds. 1975. Work Incentives and Income Guarantees: The New Jersey Negative Income Tax Experiment. Washington, D.C.: Brookings Institution.

Poirer, Dale J. 1977. "The Determinants of Home Buying." In Harold W. Watts and Albert Rees. eds. The New Jersey Income-Maintenance Experiment. Vol. 3. New York: Academic Press. Pp. 73-91.

Rice, Stuart A. 1928. Quantitative Methods in Politics. New York: Knopf.

Roistacher, Elizabeth. 1974. "Residential Mobility." In James N. Morgan. ed. Five Thousand American Families--Patterns of Economic Progress. Vol. 2. Ann Arbor: Institute for Social Research, University of Michigan.

Roistacher, Elizabeth. 1975. "Residential Mobility: Planners, Movers and Multiple Movers." In Greg J. Duncan and James N. Morgan. eds. Five Thousand American Families--Patterns of Economic Progress. Vol. 3. Ann Arbor: Institute for Social Research, University of Michigan.

Rossi, Peter H. 1981. "Residential Mobility." In Katherine Bradbury and Anthony Downs. eds. Do Housing Allowances Work? Washington, D.C.: Brookings Institution. Pp. 147-183.

Rossi, Peter H. and Katherine C. Lyall. 1976. Reforming Public Welfare: A Critique of the Negative Income Tax Experiment. New York: Russell Sage Foundation.

Schnare, Ann B. 1979. Household Mobility in Urban Homesteading Neighborhoods: Implications for Displacement. Washington, D.C.: U.S. Government Printing Office

Sewell, William H. and Robert M. Hauser. 1980. "The Wisconsin Longitudinal Study of Social and Psychological Factors in Aspirations and Achievements." In Alan C. Kerckhoff. ed. Research in Sociology of Education and Socialization. Vol. 1. Greenwich: JAI Press. Pp. 59-99.

Struyk, Raymand J. and Marc Mendick, Jr. 1981. Housing Vouchers for the Poor: Lessons from a National Experiment. Washington, D.C.: Urban Institute.

U.S. Department of Commerce, Office of Federal Statistical Policy and Standards. 1978. "Longitudinal Surveys." In A Framework for Planning U.S. Federal Statistics for the 1980's Washington, D.C.: U.S. Government Printing Office. Pp. 321-325.

U.S. Department of Housing and Urban Development. 1977. The Urban Homesteading Catalogue. 3 volumes. Washington, D.C.: U.S. Government Printing Office.

U.S. Department of Labor, Bureau of Labor Statistics. 1980. Using the Current Population Survey as a Longitudinal Data Base. Bureau of Labor Statistics Report No. 608. Washington, D.C.: U.S. Government Printing Office

Watts, Harold and Albert Rees. eds. 1977. The New Jersey Income-Maintenance Experiment. Vols. 2 and 3. New York: Academic Press.

Wise, Lauress L. and Lauri Steel. 1980. "Educational Attainment of the High School Classes of 1960 Through 1963: Findings from Project TALENT." In Alan C. Kerckhoff ed. Research in Sociology of Education and Socialization. Vol. 1. Greenwich: JAI Press. Pp. 101-126.

Wooldridge, Judith. 1977. "Housing Consumption." In Harold W. Watts and Albert Rees. eds. The New Jersey Income-Maintenance Experiment. Vol. 3. New York: Academic Press.

SIPP AND CPS LABOR FORCE CONCEPTS: A COMPARISON

by Paul Ryscavage, Bureau of the Census

Background

The Survey of Income and Program Participation (SIPP) is a new Census Bureau survey designed to give policymakers, researchers, and the public an in-depth look at the economic situation of persons and households in the United States. Its primary purpose is to collect data on the kinds and amounts of income received by persons and the extent of their participation in government income transfer programs, such as Social Security and Aid to Families with Dependent Children. The full scope of SIPP as a source of information on the well-being of our society, however, is still being realized.

One important byproduct of SIPP is information on the labor force activity of individuals. Working or not working is frequently associated with one's economic situation and also one's participation or nonparticipation in social welfare programs. An obvious illustration is the relationship between job loss and the receipt of unemployment insurance payments.

In the development of the SIPP labor force questions, an effort was made to make them conceptually similar to those in the Current Population Survey (CPS) which is the survey used to collect the Federal government's official labor force statistics. The CPS, in operation since 1940, was developed for the sole purpose of estimating the numbers of employed and unemployed persons in the country.

At the core of the CPS labor force data is the "activity concept."[1/] Basically, the concept amounts to identifying persons' activities in relation to the labor market during a specific period of time. In the CPS the period of time is one week. Persons in the adult population can then be sorted into three mutually exclusive groups depending on their activity during the week: working, not working but seeking work ,and neither working nor seeking work.

While many refinements have been made to the activity concept and the operation of the CPS through the years, the keystone of the Nation's employment and unemployment estimates--activity during a specific reference week--has not not been changed. The concept and the CPS have been reviewed periodically by Presidentially appointed commissions to insure their soundness. The most recent review was by the National Commission on Employment and Unemployment Statistics (NCEUS) in the late 1970's.[2/] Although the Commission recommended some modifications of definitions used in the CPS, it pronounced the basic activity concept as sound.[3/]

Compared to the CPS, SIPP is in its infancy. Its genesis was the Income Survey Development Program begun in the mid-1970's by the Department of Health, Education, and Welfare.[4/] Despite its newness, SIPP has great potential for not only casting light on the nature of income dynamics, but also on how labor force activity is related to it. Indeed, the NCEUS suggested there was a need "to link labor force experience with income data" so as to add a qualitative dimension to labor force statistics.[5/] SIPP data will show on a regular basis how well the labor market is providing for the economic well-being of workers and their households.

An obvious question among labor force analysts is how will the SIPP and CPS labor force data compare? Although we can't answer that question at this time because SIPP labor force data are still being processed, we can compare SIPP and CPS labor force concepts.[6/] More specifically, we can examine how the activity concept is applied in both surveys. We begin first by briefly reviewing some of the survey design characteristics of the SIPP and CPS and then compare specific SIPP and CPS labor force definitions. A concluding section of the paper discusses potential uses of SIPP labor force data.

Survey Design Characteristics of SIPP and CPS

Labor force analyses (as well as other kinds of analyses) are frequently limited because the data being analyzed come from surveys with unique survey design characteristics. For example, small sample size often creates difficulties for analysts. Three survey design features of SIPP and CPS which are important from an analytical standpoint are discussed below.

Samples. Significant differences exist in the sample size and design of SIPP and CPS. SIPP is a longitudinal panel survey comprised originally of 26,000 households located in 174 areas around the country. The sample is divided into four rotation groups and households in each group are interviewed every four months for approximately two and one-half years. The first rotation group was interviewed in October 1983, and interviews were conducted in the second, third, and fourth rotation groups in November, December, and January, respectively. This staggered sample design produces a cycle or wave of interviewing and takes four months to complete after which the rotation groups are reinterviewed in the same sequence. The Census Bureau plans to introduce another panel of approximately 20,000 households in January 1985 and another 20,000 household panel in January 1986. Consequently, SIPP's sample size will grow as panels are overlapped, increasing the reliability of the estimates.

The CPS is basically a cross-sectional surey, but it also has a longitudinal dimension.[7/] It is a much larger survey being composed of 60,000 households located in 629 areas across the country. The CPS sample is divided into eight rotation groups but unlike the staggered sample design of SIPP, all rotation groups are in operation in a single month. The longitudinal aspect of CPS results from the rotation group pattern in which a household is in the sample for four consecutive months, out for eight and then back in for four more months. This pattern allows three-quarters of the households to be the same from month-to-month and one half to be the same over the year. This is important because labor force analyses of CPS data conducted by the Bureau of Labor Statistics (BLS) concentrate on month-to-month and year-to-year changes.

Two problems for labor force analysts who use household survey data are biases resulting from the sample's design and from interview nonresponse. Rotation group bias has always been a problem in the CPS and it has received much attention over the years.[8] Theoretically, each CPS rotation group should produce the same estimates, except for random differences due to sampling variability. The estimate of unemployment from the first rotation group, however, is usually greater than the estimate based on all rotation groups.[9] (Recently, the difference has averaged about six percent.) The reason for the difference has never been isolated. Because all SIPP rotation groups in a SIPP panel have been in the sample for the same amount of time, this type of bias will not be immediately observable. It will be possible to observe after the introduction of the 1985 SIPP panel in January 1985, however, since rotation groups of different sample ages will then be in operation at the same time.

A second bias problem involves survey nonresponse--unit or total nonresponse and item nonresponse. In the CPS, the unit noninterview rate has hovered around the four to five percent mark in recent years; item nonresponse rates vary by item, but in the March CPS income questions generally have the highest nonresponse rate.[10] The Census Bureau has developed noninterview adjustments and imputation schemes for dealing with these problems. While the first panel in SIPP is less than a year old, it appears that the unit noninterview rate for the first SIPP interviews is about the same as in the CPS. (A cumulative noninterview rate will be available from SIPP as subsesquent waves of interviewing is completed.) Item nonresponse in SIPP is presently being investigated at the Census Bureau.[11] Because both labor force and income questions are asked at the same time, the quality of the SIPP labor force data may be affected.

Survey eligibility and coverage. Respondent eligibility and coverage are somewhat different in SIPP and CPS. In SIPP all household members 15 years of age and over are eligible to be interviewed and all eligible persons are interviewed if present at the time of the interview. If an eligible person is not home, a "proxy" interview is obtained from a knowledgeable person, otherwise a return visit is scheduled. In the CPS the age of eligibility is 16 years and over (data are also collected for 14 and 15 year olds); one adult household respondent may answer the questions for all household members.

Telephone interviewing is also handled differently in the two surveys. Telephone interviews in SIPP must have prior regional office approval, except in the case of information not obtained in the course of the interview. In the the CPS, telephone interviews are permitted in the second, third, fourth, sixth, seventh, and eighth month in which the households are in sample.

Another difference concerns the treatment of the Armed Forces. In the monthly CPS, members of the Armed Forces living in households are not eligible for interview. In SIPP, however, such individuals are interviewed as long as they are stationed in the area and usually reside at the address visited. (Both surveys exclude inmates of institutions, such as persons in prisons or convalescent homes.)

Lastly, and most significant for many analyses, members of households in the SIPP sample who move between interviews are followed and further interviews attempted. Sample persons, however, are not followed when they have been institutionalized, become a member of the Armed Forces, move outside the United States, or move more than 100 miles from a SIPP sampling area. In the CPS, movers are not followed and this has been a constraint on many longitudinal labor force analyses.[12]

Reference periods. A fundamental difference between SIPP and CPS--one that will probably account for differences in labor force estimates between surveys -- is the length of the reference period. CPS interviews are conducted in all rotation groups each month in the week containing the 19th and all questions about labor force activity are asked in reference to the previous week which contains the 12th, the survey week. (As will be discussed, this one week reference period is extended to four weeks in the case of jobseeking.) Depending on the respondent's answers to the questions, household members are classified into one of three mutually exclusive groups, employed, unemployed, or not in the labor force.

SIPP interviews are conducted in one of the four rotation groups each month during the first two weeks of the month. The labor force, income, and program participation questions relate to the four previous months. Indeed, the labor force questions actually refer to individual weeks during the four month period. During this time a person could have worked, looked for work, and been outside the labor force at different times. In other words, labor force classification in SIPP is not necessarily mutually exclusive as it is in the CPS.[13]

Recall problems are potentially a greater problem in SIPP than in the monthly CPS since respondents are recalling activities over a much longer period. For persons with a marginal attachment to the labor market, for example, teenagers, it may be very difficult to remember job market activities. Despite the long recall period in SIPP, it is not inordinately long. In the supplement to the March CPS, persons are asked about their labor market activities in the previous calendar year--a reference period extending back 3 to 15 months.[14] The annual work experience statistics have been published by the BLS and Census Bureau for years.

SIPP and CPS Labor Force Definitions

Because the reference periods in SIPP and CPS are of different lengths, the activity concept is applied differently in both surveys. In the CPS, persons are asked a specific activity-type question relating to the week containing the 12th of the month. In sorting out the possible labor market-related activities into mutually exclusive groups, a priority scheme is necessary since some individuals may have been involved in more than one activity. The first or highest priority is assigned to working. As long as a person worked for pay or profit for one hour or more (or 15 hours or more without pay in a family operated business), the person is considered employed even though he or she may also have looked for work or gone to school or done something else during that week.

The next highest priority is given to those persons who had a job during the survey week, but were temporarily absent from it. Although this relaxes the activity concept slightly, it permits a more accurate counting of the numbers of persons with actual job commitments. These persons may have been on strike or ill or on vacation or absent for some other personal reason, but since they had a job to return to they are classified as employed.

The third priority is assigned to persons whose activity was looking for work. If a person in the survey week neither worked nor had a job but looked for one at some time within the past four weeks (and was currently available to take one) he or she is considered unemployed. Once again the activity concept is relaxed to cover persons who may not have looked for work continuously because they were waiting to be recalled from layoff or were waiting to start a new wage or salary job within 30 days. These persons too would be classified as unemployed. All other individuals not fitting into this classification scheme are considered to be not in the labor force. Accordingly, the Nation's civilian noninstitutional population age 16 and over can be sorted into the familiar labor force categories shown below.

CPS Labor Force Categories

Civilian noninstitutional population age 16 and over..
  Civilian labor force ....
    Employed ....
    Unemployed ....
  Not in the labor force ....

In SIPP persons are not asked a specific activity question relating to the previous four months because it is very possible (even more so than in the CPS) that an individual may have worked, looked for work, or done something else during the period. Instead, the initial question on the SIPP questionnaire concerns whether or not an individual had a job or business at any time during the previous four months. In other words, the activity concept is tied to an individual's either having or not having a job in the reference period. For those who had jobs, subsequent questions are asked about how long persons had their jobs in the reference period, whether they had been absent from them and why, and whether they looked for work or were on layoff when they did not have jobs. For persons who did not have jobs during the entire period, questions are asked if they looked for work or had been on layoff and if so, for how long.

Unlike the CPS where a priority scheme is required to classify individuals into mutually exclusive labor force categories, in SIPP individuals may have experienced more than one labor force status in the four month reference period.

For example, a person may have had a job for the entire period but was temporarily laid off for one month of the period; or a person may have had a job for two months and then quit to look for another one; or a person may have been outside the labor force for a month, looked for a job for a month and then, having found one, worked for two months. Consequently, the SIPP labor force categories shown below reflect multiple labor force statuses.

SIPP Labor Force Categories

Noninstitutional population age 16 and over....
  Persons with some labor force activity in period ....
    With a job the entire period ....
      Worked all weeks ....
      Missed some weeks ....
        Spent time on layoff ....
    With a job during part of the period .....
      Spent time looking for work or on layoff
    No job during the entire period ....
      Spent time looking for work or on layoff entire period ....
      Spent time looking for work or on layoff some weeks ....
  Persons with no labor force activity in period ....

A closer look at specific SIPP and CPS labor definitions is presented below. (The definitions are discussed under headings common in everyday usage and should not be construed as CPS-specific labor force terminology.)

Employment. In both surveys, employment is generally defined as working at a job or business for pay or profit at some time in the reference period. A job is considered to be an arrangement for regular work for pay where payment is in cash wages or salaries, at piece rates, in tips, by commission, or in-kind (meals, living quarters, supplies received). A business is defined as an activity which involves the use of machinery or equipment in which money has been invested or an activity requiring an office or "place of business" or an activity which requires advertising. Payment may be in the form of profits or fees. Both surveys also consider persons to be employed when they have been absent from their jobs because of illness, vacation, bad weather, labor dispute, and various personal reasons. Unpaid family work is considered employment when it contributes to the operation of a farm or business run by a member of the same household. In the CPS, unpaid family work must have lasted for 15 hours or more during the reference week, but in SIPP there is no hours restriction.

Unemployment. The definitions of unemployment in both surveys are also very similar. In CPS, persons must have been without a job during the reference week and in SIPP they must have been without a job for all or part of the reference period; in addition, they must have been available for work, and taken some specific jobseeking activity. Job seeking activity in CPS may have occurred anytime in the previous four weeks, while in SIPP it may have occurred any time during the four months. If, in either survey, job seeking occurred when the person was working, working would take precedence and the person would be considered employed.

Two exceptions to the above rule must be noted. The first is the case of the person who has a job but was laid off and the second is the person who is to begin a new wage or salary job within 30 days. Both persons are considered unemployed. In the CPS these persons must have been available for work, but in SIPP no availability test is applied.

Because the CPS is basically a labor force survey, it collects more information about the spell of unemployment than SIPP. For example, CPS gathers information on reasons for unemployment whereas SIPP does not. One can tell from CPS data whether a jobless person has become unemployed because of job loss, such as a layoff; quitting a job to search for another; entering the labor force for the first time; or re-entering the labor force. In SIPP, the only group for whom the reason is known for being unemployed are those persons who report themselves as having jobs from which they are absent because of layoff. The CPS also asks about the method of job search and how long one has been searching or on layoff.

Labor force. The civilian labor force in the CPS is derived by adding the number of persons classified as employed during the reference week to the number who were classified as looking for work or on layoff. The "total" labor force is derived by adding to the civilian labor force an independent estimate of the Armed Forces stationed in the United States.

The labor force in SIPP (which includes members of the Armed Forces living in households but not in installations of the Armed Forces) is is referred to as "Persons with some labor force activity." This represents the sum of persons who, during the four month reference period, may have been --

- employed during all weeks,
- unemployed during all weeks,
- employed and unemployed during all weeks,
- employed and outside the labor force during all weeks,
- unemployed and outside the labor force during all weeks,
- and employed, unemployed, and outside the labor force during all weeks.

In other words, any one with some contact with the labor market in the four month reference period.

Unemployment rate. The unemployment rate from the CPS is one of the most well known statistics in the Nation. It is derived by dividing the number of unemployed persons by the civilian labor force (or total labor force). In SIPP a similar rate, or proportion, could be calculated. Unlike the CPS unemployment rate definition, however, the numerator in the SIPP definition is composed of persons who may have been

- unemployed during all weeks,
- employed and unemployed during all weeks,
- unemployed and outside the labor force during all weeks,
- and employed, unemployed, and outside the labor force during all weeks.

In other words, the numerator is composed of "Persons with some unemployment." Dividing the sum of these groups by persons with some labor force activity--the denominator--will yield the proportion, or percentage, with some unemployment.15/

Not in the labor force. In both the CPS and SIPP, persons who have had no association with the job market during the reference period (in SIPP, for all or part of the reference period) are considered outside the labor force. The CPS further identifies their major activity as in school, keeping house, unable to work, and so on. This is not done in SIPP.

The CPS inquires in the fourth and eighth rotation groups about previous work experience, intentions to seek work again, desire for a job, and reasons for not looking. This makes it possible to estimate the number of "discouraged workers." Discouraged workers in the CPS are defined as persons who want a job but are not seeking work currently because: 1) they believe no work is available in their line of work or area; 2) they could not find any work; 3) they lack the necessary schooling or training, skills, or expeience; 4) employers think they are too young or old; and 5) they have other personal handicaps in finding a job, such as transportation problems.

An effort is made to identify discouraged workers in SIPP also, even though it is difficult to recall a state of mind. For those persons who did not work or look for work in at least part of the four month reference period but said they wanted a job and could have taken one, a question is asked as to why they were not looking. The reasons for not looking are very similar to those in the CPS questionnaire.

Hours of work. In the CPS a question is asked about the number of hours some one worked during the reference week at all jobs. This question is asked of all rotation groups and includes workers who have more than one job. In addition, in two of the eight rotation groups a question is asked about the hours "usually" worked at the worker's main job. This information is part of the CPS data collected on workers' earnings.

A similar set of questions is found in SIPP. Everyone who worked is asked about their usual weekly hours on all jobs during the four month period. Subsequent questions inquire about usual weekly hours for the primary job and any others.

Full-time and part-time employment. Full-time employment in both surveys is defined as employment of 35 hours a week or more while part-time employment is anything less than 35 hours. Both surveys seek the reasons for part-time employment, that is, whether it was due to economic reasons or other factors. Economic reasons include slack work, material shortages, repairs to plant or equipment, start or termination of a job during the week, and the inability to find full-time work. "Other" reasons include labor disputes, bad weather, one's own illness, vacation, keeping house, no desire for full-time work, and full time worker during only part of the season. In the SIPP questionnaire the reasons for part-time employment are not as numerous but it is still possible to identify some economic reasons for part-time employment.

Uses of SIPP Labor Force Data

SIPP was designed primarily as an income survey and the data from it will be used to address issues related to income security and social welfare programs. With the inclusion of questions on labor force activity, however, this survey has potential for labor force analysis and topics related to it. In addition, because of SIPP's sample design both cross-sectional and longitudinal data can be obtained from the survey providing analysts with more flexibility in their analyses. For example, it is possible to calculate monthly averages of the labor force data from SIPP waves since labor force activity is tracked (week-by-week) over a four month period; on the other hand, by linking all the SIPP waves it is possible to follow the labor force activity of individuals over two and one-half years.

While the CPS will continue to be the primary source of information on the country's labor supply and the current unemployment situation, SIPP labor force data will complement the basic CPS information in many ways. The following is a discussion of some of the applications of SIPP cross-sectional and longitudinal labor force data.

Labor market related economic hardship. For many years economists have tried to measure the economic hardship caused by labor market problems, whether they be demand oriented (unemployment due to insufficient jobs) or supply oriented (low wages because of insufficient skills and education). The economic literature contains many references to subemployment indices, employment and earnings inadequacy indices, and labor market hardship measures of one variety or another.16/ The NCEUS in 1979 examined this subject and recommended that the BLS publish an annual report "... containing measures of different types of labor market related economic hardship resulting from low wages, unemployment, and insufficient participation in the labor force."17/ Using data from the March CPS, the BLS has produced such reports but they are not as comprehensive as they might be because of data limitations (for example, neither the hourly earnings for part-year workers nor the problems of discouraged workers are discussed.)18/

SIPP labor force and income data should be able to fill the gap. For example, one cross-sectional table specification might show employment problems incurred by individuals cross-classified by their position in the household income distribution. Problems of unemployment, low hourly wages (below the Federal minimum), discouragement, and involuntary part-time employment could be isolated to help in formulating applicable policies. This information, in combination with income information, is available on a current basis only from SIPP.

Labor mobility and turnover. Given the longitudinal feature of SIPP's sample design, not only can the income flows and program participation activities of individuals be monitored for two and one-half years (and periods of shorter duration), but so can their labor force activities. At the time of each SIPP interview, information is obtained on the labor force activity of each household member age 15 and older during the prior four months. Any changes in labor force status during this period are reflected in the data. Stitching together the data collected in each of the eight or nine interviews will provide data users with a profile of labor market activity for a two and one-half year period.

One change in labor force status that labor economists have been interested in recently is the one which occurs after a spell of unemployment. Some have argued that many outcomes of spells of unemployment are withdrawals from the labor force and not reemployment. For example, two economists, using CPS gross flow data, estimated that 45 to 50 percent of all unemployment spells end by labor force withdrawal.19/ Other economists have argued that the relative shortness of the average unemployment duration shows that persons can quite easily find their usual type of employment in a short period of time.20/ With SIPP labor force data it will be possible to identify job terminations, observe spells of unemployment, and determine not only their durations, but their outcomes.

SIPP labor force data should also be useful in calculating rates of job separation and accession. The measurement of the amount of job separation and accession is an important element in understanding our basic employment and unemployment statistics. Since the discontinuance of the BLS's labor turnover series, researchers have been hard pressed to find other data sources which would shed light on the dynamics of the labor market.21/ While it will not be possible from the SIPP labor force data to identify the precise nature of the separation (layoff, quit, discharge) or accession (new hire, recall), aggregate separation and accession rates could be calculated. These rates could be monitored over the business cycle.

Summary

SIPP is principally an income survey, but it contains questions on labor force activity as well. SIPP labor force data will supplement the labor force information from the CPS, the Federal government's official source of labor force statistics. Like the CPS, SIPP uses an activity concept for sorting the Nation's population into those persons involved in the job market from those who are not. A major difference between the two surveys is the length of the reference periods; in the CPS it is one week and in SIPP it is four months. The different length of time for which labor market activities are surveyed will be an important factor in SIPP and CPS labor force comparisons. Nevertheless, while the CPS will continue to tell us how many persons are employed and unemployed, SIPP will be able to tell us how well the labor market is providing for these workers and their households.

NOTE: In addition to the references footnoted, the SIPP Interviewer's Manual and CPS Interviewers Reference Manual were used in the preparation of this paper.

Footnotes

1/ For those interested in the origins of the activity concept, see John N. Webb, "Concepts Used In Unemployment Surveys," Journal of the American Statistical Association, March 1939, pp. 49-61. For a history of the CPS, see John E. Bregger, "The Current Population Survey: a historical perspective and BLS' role," Monthly Labor Review, June, 1984, pp. 8-14.

2/ See Counting the Labor Force, National Commission on Employment and Unemployment Statistics, (U.S. Government Printing Office) Labor Day, 1979. An earlier review was made in the early 1960's. See Measuring Employment and Unemployment, President's Committee to Appraise Employment and Unemployment Statistics, (U.S. Government Printing Office) 1962.

3/ See Counting the Labor Force. p. 2.

4/ See Martynas A. Ycas and Charles Lininger, "The Income Survey Development Program: Design Features and Initial Findings," Social Security Bulletin, November 1981, pp. 13-19.

5/ See Counting the Labor Force. p. 1.

6/ Labor force estimates from the ISDP were compared to the CPS by Bruce Klein. He found that employment estimates in ISDP were slightly higher than in CPS and unemployment estimates in ISDP were considerably lower than in CPS. One reason for the latter was that persons on layoff were not counted in the ISDP. See Bruce W. Klein, "Comparing Labor Force Measures in ISDP with CPS," Technical, Conceptual, and Administrative Lessons of the Income Survey Development Program (ISDP), Papers presented at a Conference (Martin H. David, ed.), Social Science Research Council, New York, New York, 1983, pp. 229-239.

7/ For a discussion of the longitudinal nature of the CPS see, Using the Current Population Survey as a Longitudinal Data Base, Report 608, Bureau of Labor Statistics, August 1980.

8/ See Barbara Bailar, "The Effects of Rotation Group Bias on Estimates from Panel Surveys, "Journal of the American Statistical Association, Vol. 70, March 1975.

9/ See the Current Population Survey: Design and Methodology, Technical Paper 40, Bureau of the Census, January 1978, p. 83.

10/ Ibid., p. 87.

11/ See John Coder and Angela Feldman, "Early Indications of Item Nonresponse on the Survey of Income and Program Participation," a paper to be presented at the 1984 Joint Statistical Meetings, Philadelphia, Pa., August 1984.

12/ See, for example, Francis W. Horvath, "Tracking Individual Earnings Mobility With the Current Population Survey," Monthly Labor Review, May 1980, pp. 43-46.

13/ A similar situation prevails in the March supplement to the CPS where persons are asked about their work experience in the previous calendar year.

14/ The retrospective bias in the March CPS work experience data has been the topic of research in recent years. For example, see Richard D. Morgenstern and Nancy S. Barrett. "The Retrospective Bias in Unemployment Reporting by Sex, Race, and Age," Journal of the American Statistical Association, June 1974, pp. 355-357. For more recent research see, Francis W. Horvath, "Forgotten Unemployment: Recall Bias in Retrospective Data," Monthly Labor Review, March 1982, pp. 40-43.

15/ The Bureau of Labor Statistics calculates a similar rate from the annual work experience data collected in the supplement to the March CPS. It is referred to as "the percent with unemployment."

16/ For example, see Herman P. Miller, "Subemployment in Poverty Areas of Large U.S. Cities," Monthly Labor Review, October 1973, pp. 10-17; Sar A. Levitan and Robert Taggart, Employment and Earnings Inadequacy: A New Social Indicator, (Baltimore: The Johns Hopkins University Press; 1974); T. Vietorisz, R. Mier, and J. Giblin, "Subemployment: Exclusion and Inadequacy Indexes," Monthly Labor Review, May 1975, pp. 3-12; and Robert Taggart, Hardship --The Welfare Consequences of Labor Market Problems: A Policy Discussion Paper, The W.E. Upjohn Institute for Employment Research, 1982.

17/ See Counting the Labor Force, p. 60.

18/ The latest BLS report is Linking Employment Problems to Economic Status, Bulletin 2201, U.S. Department of Labor, Bureau of Labor Statistics, June 1984.

19/ See Kim Clark and Lawrence Summers, "Labor Market Dynamics and Unemployment: A Reconsideration," Brookings Papers on Economic Activity, No. 1, 1979, pp. 13-72.

20/ See Martin Feldstein, "The Importance of Temporary Layoffs: An Empirical Analysis," Brookings Papers on Economic Activity No. 3, 1975, pp. 725-744.

21/ For a statement of the need for labor turnover data, see Robert E. Hall and David M. Lilien, "The Measurement and Significance of Labor Turnover," in Counting the Labor Force, Appendix Vol. 1 (Concepts and Data Needs), National Commission on Employment and Unemployment Statistics (Washington D.C., 1979), pp.577-600. For an example of separation data created from the CPS see S. Haber, E. Lamas, and G. Green, "A New Method for Estimating Job Separations by Sex and Race, Monthly Labor Review, June 1983, pp. 20-27, and Allan Eck, "New Occupational Data Improve Replacement Estimates," Monthly Labor Review, March 1984, pp. 3-10.

MATCHING ECONOMIC DATA TO THE SURVEY OF INCOME AND PROGRAM PARTICUPATION: A PILOT STUDY

Sheldon E. Haber, The George Washington University
Paul M. Ryscavage, Douglas K. Sater, Victor M. Valdisera, The Bureau of the Census

The new Survey of Income and Program Participation (SIPP) will undoubtedly become a major source of data on a wide variety of aspects of the well-being of our nation's households, families, and individuals. The very richness of SIPP suggests the desirability of augmenting it with micro-level establishment and enterprise data from the economic censuses and other data files maintained by the Bureau of the Census, since the marginal cost of merging these data with SIPP is relatively small and the potential gain in knowledge is very large. One area where the payoff relative to cost of enhancing SIPP is sure to be substantial and significant is that pertaining to the behavior of labor markets.

A list of some of the areas in which a SIPP-economic data file can yield new insights includes the following topics:

The relationship between capital and wage rates
Labor mobility
Low wage workers and low wage firms
Measuring the effects of minimum wage legislation
Structural unemployment
Identifying high tech workers and high tech firms
Implications of the transition from a goods to a service economy
Unions and the labor market
The substitutability of capital and labor
Productivity analysis

Besides the substantive knowledge to be gained by merging SIPP demographic and economic data, there are externalities associated with merging these data sets. First, it will be possible to verify the accuracy of the size of firm estimates given by respondents in survey data. An additional, indirect benefit of linking SIPP and economic data stems from the fact that the former is a representative sample of the working population. Matching on work place will yield a stratified sample of firms where the probability of selection is inversely proportional to firm size. By weighting the number of firms in each size group, estimates for the entire population of firms can be derived. The sample of employers would be contained in a single data set versus the diversity of data sets in which the economic data are now found, with the same format across employers. These advantages plus the manageable size of the sample should provide valuable insights into the structure of production within and across sectors of the economy at a point of time and over time.

1. SIPP and the Economic Data Files In merging demographic and economic data, it is necessary to know the information contained in the various files to be linked and how each file is constructed. In this section, we briefly describe four data sets which might be incorporated into a SIPP-economic data file.

SIPP contains demographic and program related data. Economic data are found in the Standard Statistical Establishment List (SSEL), the Longitudinal Establishment Data (LED) file, and the enterprise statistics (ES). The SSEL covers all establishments and companies with employees and yields current information on employment and payroll. The LED, as its name implies, contains longitudinal data but is restricted to manufacturing establishments. The ES, on the other hand, covers companies in the construction, mineral, manufacturing, wholesale trade, and retail trade industries, and most service industries.

The SSEL is a complete directory of establishments in single and multi-establishment enterprises with one or more employees, irrespective of industry. The SSEL links parent companies, subsidiaries, and their establishments. It contains information on approximately 4.7 million enterprises and 5.7 million establishments.

The importance of the SSEL is that it is a current file containing a complete list of establishments and companies with paid employees. While the SSEL contains a narrow range of economic data, these data impart valuable information. For example, the SSEL contains the address of the physical location of establishments which is useful for merging the demographic and economic data, since it is a primary link in identifying an individual's place of work. Employment and payroll figures yield an estimate of average annual earnings, thereby indicating whether an employer is a low or high wage employer. [1/] Sales and employment figures provide a proxy measure of productivity. Operational status information can be utilized to identify establishments which have become inactive. Additionally, the SSEL contains longitudinal information. Currently, establishment and company data are carried for two years in the SSEL.

The LED is a longitudinal micro-data base containing data at the establishment level from the Annual Survey of Manufactures and the Census of Manufactures.

The LED provides a much broader range of information about establishments than the SSEL. For each manufacturing establishment, value added per production worker, which is a measure of labor productivity, can be calculated. For the larger establishments with 250 or more workers, information is available on depreciable assets and rented machinery so that capital/labor ratios can be computed. Also, a better measure of labor compensation, including fringe benefits, can be obtained.

Like the Census of Manufacturers, the enterprise statistics (ES) are collected every five years. These data cover enterprises whose primary activity is in an in-scope industry. For each enterprise, the data are consolidated over all operating units. The information contained in the ES is similar to that in the Census of Manufactures, except that fringe benefit, asset, and related data are available only for companies with 500 or more workers.

2. Some Applications of Micro-Demographic and Economic Data In this section, three applications of a SIPP-economic data file are discussed to illustrate how this data set can help explain policy issued relating to earnings and employment.

A. Low Wage Workers and Low Wage Firms While survey data such as the CPS provide insights into the characteristics of low wage workers, they provide no information about low wage firms. Despite this lack of information, an *a priori* description of low and high wage firms can be formulated. All else being the same, low wage firms will be labor intensive and, hence, tend to be smaller than high wage firms. And because recruitment and hiring cost relative to the level of wages will tend to be high, low wage firms will also advertise less for labor and employ fewer screening devices to weed out suitable workers; thus, their work force will be of lesser quality than their high wage counterparts. Less qualified workers, on the other hand, e.g., younger workers and those who are less educated, will be attracted to low wage firms because their marginal product is less than that required to gain employment in high wage firms. More generally, workers with given characteristics and tastes sort themselves among firms with similar requirements for labor.

Corresponding to the greater prevalence of low quality workers in low wage firms, one might expect that in these firms (vis-a-vis high wage firms) a higher proportion of capital expenditures is for used rather than new machinery and equipment; likewise, the proportion of depreciable assets retired each year is likely to be smaller in such firms. Furthermore, given that labor is of lesser quality and capital is of an older vintage, it would not be surprising if value added per worker were relatively low in low wage firms.

Other characteristics are more easily seen by focusing on high wage firms. To the extent that high wage firms are capital intensive, their need for trained workers is likely to be greater than that of low wage firms. Capital intensiveness suggests greater use of resources to monitor output; hence, a higher proportion of the work force may be needed in supervisory positions. To reduce turnover, which disrupts the production process, high wage firms will substitute future benefits in the form of pensions for current benefits in the form of wages. A SIPP-economic data file would permit verification of these hypotheses.

Information about low and high paying firms is important for another reason besides the light it sheds on how production is organized in these two types of firms. Since low paying firms are a source of employment for workers with relatively low productivity, it is of some interest to inquire into the extent to which low pay among workers is attributable to their employment in such firms. In approaching the question of why some workers are paid less than others in this manner, low wage employers can be viewed as providing employment opportunities with attendant low earnings, not because they descriminate against certain groups of individuals, but because the production processes that are most efficient for their mode of operation do not require high quality labor and, furthermore, they inhibit their paying high wages.

A procedure for verifying this view would be to sector firms according to whether they are low paying or high paying. With this sectoring of firms, one would expect, as indicated above, that the mix of workers and capital is dissimilar between the two sectors. Assuming this is so, to what extent are differences in individual earnings in low and high paying firms due to the characteristics of the workers and capital employed in each type of firm? Also, to what extent are workers with similar characteristics renumerated in the same way in each type of firm? The answers to these questions can be obtained from a SIPP-economic data file.

B. Structural Unemployment An issue of long standing is what happens to workers who are displaced from their job as a result of structural disequilibria. How long do they remain unemployed vis-a-vis other workers who separate from an employer? What sources of income, including cash and noncash government transfers, do they draw on when they are unable to find work? When they find a job, how do earnings in the new job compare to earnings in the old one? If there is an earnings loss, how much of this loss is recouped, say, after 2 years?

A major problem in answering these questions is that workers do not know if they are structurally unemployed. One way of identifying such workers is to ascertain what has happened to the firms in which they were last employed.

If the firm has undergone a substantial decline in employment or has closed down for a a relatively long period of time, say, longer than the typical recession, one may presume that it has undergone a shock which is typical of the shocks experienced by firms subject to structural disequilibria. It also can be presumed that the employees of these firms experience the aftereffects of such shocks.

A SIPP-economic data file would enable one to determine the extent to which firms are subject to severe, long-term shocks as evidenced by plant closures and substantial reductions in employment, and how such shocks affect their work force.

C. High Tech Workers and High Tech Firms Despite the importance of new technologies for improving productivity, there is no widely accepted definition of a high tech industry. Based on a definition which includes industries with a ratio of technology-oriented workers [2/] to all workers of at least 1.5 times the industry-wide average, Riche, Hecker, and Burgan (1983) estimate that 13.4 percent of all wage and salary workers were employed in high tech industries in 1982.

High tech industries have been cited as having a large group of high and low wage workers whereas other industries are comprised of workers who are concentrated in the middle of the earnings distribution. While it is useful to know how workers in high tech and other industries differ and the differential growth of employment in the two kinds of industries, it is equally important to know the characteristics which differentiate high tech from other firms and the differential in the rate of growth of the two types of firms.

As is self-evident, not all firms in high tech industries utilize the latest technology, and new techniques of production are utilized by firms in industries besides those labeled as high tech. One approach to distinguishing between the two types of firms would be to compare the characteristics of the industries denoted on a priori grounds as high tech with other industries and then to use this information to identify high tech firms. To illustrate this approach, assume that the a priori criterion used to denote high tech industries is the one noted above, namely, that the ratio of high tech to all workers in a given industry to the similar ratio for all industries is higher than some minimum value. Assume also that the high tech industries exhibit high values of the following ratios: capital expenditures for new computers to allcapital expenditures, capital expenditures to asset value, and capital to labor. Given a set of characteristics which permit the bifurcation of industries, the multivariate technique of cluster analysis can then be applied to identify high tech firms within both high tech and other industries.

The outcome of the cluster analysis is a partitioning of firms into categories, i.e., high tech and nonhigh tech firms, as determined by the data, where each cluster of firms represents a homogeneous set of observations. An advantage of applying the aforementioned two-stage procedure using a SIPP-economic data file is that it provides an independent test of how well the procedure works. For if the approach is successful, the proportion of workers who are technology-oriented among the firms classified as high tech (taken as a group) will be higher than the similar proportion for firms classified as nonhigh tech (again, taken as a group), and the difference in proportions will be greater than the corresponding difference when industries are classified as high tech and nonhigh tech.

Having identified high tech firms, in contrast to high tech industries, insights can then be obtained as to how production processes in these firms differ from their nonhigh tech counterparts. At the same time, it will enable one to better define high tech occupations and how workers in these (and other) occupations in high tech firms differ from similar workers in nonhigh tech firms.

3. The Pilot Study A principal part of the pilot study is designed to assess the availability, sources, coverage, and content of the various economic data files maintained by the Bureau of the Census and to explore study areas and issues to which a data set combining micro-worker and firm data would be applied. In the course of this study, specific demographic and economic variables have been identified which should be incorporated into such a data set. Additionally, it was anticipated that methodological problems inherent in this undertaking would be revealed; indeed, this has been the case.

A second phase of the pilot study is to investigate the efficiency of four alternative methods of identifying an individual's employer. Each method is based on different information for searching the SSEL and identifying the employer's census file number (CFN). The first utilizes information on employer name, the state of residence and/or zip code of the employee, and census industry code. The same information is used in the second method; however, additional reference materials, e.g., 1980 Census Company Name and Place of Work lists, Dun and Bradstreet reference books, Standard and Poor directories, and telephone books, will be used to obtain the exact address of an individual's employer. The third method uses the employer's name and exact address if known. In the last method, if the employer's identification number (EIN) is known, it is used in conjunction with the information available in the first three methods to identify the employer's CFN. For each method, match rates and cost information will be developed for a small sample of workers.

A third phase of the study is the construction of a pilot SIPP-economic data file in which the SIPP portion of the file would be restricted to full-time workers in large manufacturing establishments; the source of the economic data would be the LED. The objective in this phase is to calculate match rates between workers in SIPP and their establishments in the LED.

Given the importance of the wage determination process, one of the areas noted above, e.g., low wage workers and low wage firms, would be studied when the pilot work file is completed. Demonstration of the utility of this research endeavor in terms of its contribution to the economic literature would constitute the final phase of the pilot study.

4. Methodological Problems in Matching Demographic and Economic Data In this section, attention is focused on two methodological problems. The first problem deals with procedures for tying workers to their establishment and company. The second relates to the estimation of data, in particular, asset and fringe benefit data, which although available for large establishments and companies, are generally not collected for small ones.

Central to the creation of a SIPP-economic data file is the ability to determine the establishment and/or company in which a person is employed. The most promising and least expensive way of doing this is to match on firm name and physical address of an individual's place of work. This information will be available in SIPP and is available in the SSEL. Although the physical address is not necessary for identification of an individual's work place, its availability greatly facilitates such identification since a firm may have more than one establishment in a local area.

For employers with only one establishment in an area, the firm name and employee's address will typically be sufficient to determine where a person is employed. As noted, for companies with more than one establishment in an area, the firm name and employer's address should be sufficient to identify the place of work. If an employer has more than one establishment in an area and the place of work cannot be determined using the employer's physical address, or no address is available, other information in SIPP can be utilized. Firm name, respondent's address, census industry code, and respondent's estimate of size of establishment and company can be used to identify a person's work place. For example, it is unlikely that a firm manufacturing bottles will have more than one large plant in a local area.

Another aid in identifying an individual's work place is the EIN. While a company may have a number of establishments in a local area, its subsidiaries, when identified by their own EIN, may have only one establishment in the area. Thus, the EIN of the employer for whom an individual works can be sufficient to uniquely determine the establishment in which that person is employed.

In the event that an unique work place cannot be determined for a multi-establishment firm, the employer's characteristics can be imputed. Data from the SSEL on number of employees and payroll can be averaged over a company's establishments in a local area. From the ES file, average values can be computed for variables not contained in the SSEL. For example, the average capital/labor ratio for a company with a chain of fast-food stores can be used as an estimate of the capita/labor ratio for each store in the chain.

Where it is not possible to identify a worker's firm by name in the SSEL, imputations can be made by averaging over establishments in the same local area and with the same census industry code as that of the given employer. Additionally, it may be possible to refine the imputation process by considering information contained in SIPP, e.g., the size of establishment in which an individual works and whether the firm has one or more than one establishment.

As indicated, information on assets and fringe benefits is not generally available for small establishments, but such information is available for a large sample of small establishments in manufacturing. Despite the fact that asset information is not collected for many of the firms in which individuals work, the use of an economic model, including industry, firm size, and other variables, may enable one to obtain reasonably accurate estimates of capital for small establishments.

Economic theory sugests a number of relationships which influence the amount of capital that a firm employs in its production process. In particular, since capital intensity varies with establishment size in closely related industries, it seems reasonable to assume that information about the number of employees in an establishment can be used to further refine estimates of its capital assets. All else being the same, one would expect the smaller an establishment, the lower would be its capital/labor ratio. 3/ Additionally, holding everything else constant, including establishment size, low wage establishments will substitute labor for capital in order to economize on the use of the relatively expensive factor, i.e., capital. Thus, low wage establishments will tend to have a lower capital/labor ratio than high wage establishments.

Even among establishments of the same size whose wage rate is also the same, one would expect a lower capital/labor ratio the higher the proportion of production workers among all workers. When the proportion of production workers among all workers is high, or conversely, when the percentage of workers who supervise production is low, this comes about because a firm has few assets, relative to labor, to monitor. Additional relationships between assets and other variables may exist. For example, it may be that newer establishments in an industry are more capital intensive than older ones; likewise, regional variations in entrepreneurial ability may give rise to corresponding variations in capital intensity.

Besides economic relationships, engineering relationships also may be useful in estimating capital intensity. For example, it is plausible that an establishment's capital/labor ratio is positively related to purchased electricity per employee.

Finally, an economic model can also be utilized to estimate fringe benefits for small establishments and small companies. It is plausible to assume that fringe benefits in a firm are related to its size, average wage level, legal form of organization, industry, and region where it is located. With a SIPP-economic data file more refined estimates of fringe benefits per employee can then be obtained by taking account of the percentage of employees who are covered by life and medical insurance and a private pension plan in a given group of firms, say, (small) high paying establishments in manufacturing. Given this information, the average value of these benefits per covered and noncovered worker can be calculated for each establishment in the group.

An economic model can also be developed to estimate fringe benefits per covered and noncovered worker in small establishments and companies. With appropriate information in SIPP, such estimates could provide a basis for imputing an important component of private noncash benefits to individual workers. Although it should be evident from the discussion of this paper, this last illustration is indicative of the benefits to be derived from a SIPP-economic data file.

FOOTNOTES

1/ The data referenced in this section as well as the remainder of the paper are available in the economic and (where applicable) in the SIPP data collected by the Bureau of the Census.

2/ Defined as engineers, life and physical scientists, mathematical scientists, engineering and science technicians, and computer specialists.

3/ An estimate of a firm's assets can then be obtained by multiplying the capital/labor ratio estimate by the number of workers in its employ.

REFERENCES

Riche, Richard; Hecker, Danial; and Burgan, John, "High Technology Today and Tomorrow: A Small Slice of the Employment Pie," Monthly Labor Review, November 1983, pp. 50-58.

COMMENTS

BY MARTIN DAVID, University of Wisconsin--Madison

1. Some Problems Common to All Three Papers

Selection

The SIPP is subject to a different kind of selectivity than earlier cross section surveys with which we are familiar. Understanding that selectivity is an important part of the research that must be carried out on SIPP before we can be comfortable with its results. The selectivity that affects migration and labor force measurement comes from the following rules, attrition, rules for handling first wave household non-response, and treatment of the sample universe. The selectivity that affects matching comes largely from the design of economic data samples. Each of these problems will be discussed below.

Truncation of Measures

A number of devices in the questionnaire result in logical exclusion of certain kinds of data. The respondents report detail only on two employers; student status is revealed only for some persons; income is fully reported only for some children. This truncation is highly pertinent to migration and labor force measures; less important for matching.

Attitudinal measures

The SIPP is almost entirely lacking in attitudinal measures. We need to ask to what degree subjective measures are important for the three areas covered by the papers above.

Left censoring

Although the SIPP records much current information, the duration of status at Wave 1 is conspicuously lacking. We do not know for how long the respondent was (un)employed prior to the first interview; for how long (s)he has held the current position; or for how long (s)he has resided in the current residence. Each of these items must be known to understand episodes or events in the lives of real people and will be important conditioning information for analyses of SIPP.

2. Issues pertaining to Migration

Dahmann correctly stresses the unique potentialities of SIPP for migration statistics. Unlike earlier devices for measuring migration, the SIPP gives high resolution to the timing of migration events, and associated economic and demographic circumstances before and after the migration event.

By contrast retrospective reports of migration are severely biased, because measures of migration are only obtained for those individuals who remain in the household universe. Persons who emigrate and persons who die are excluded. If their migration experiences are different from the remaining household population, retrospective measures are distorted.

Unfortunately SIPP is not a perfect instrument for recording migration because the sample systematically censors certain types of migration. Firstly, the sample is not updated with new addresses until after 12 months of operation. Thus immigrants to the household population are not represented, unless they move into existing households. This contrasts with the design of the CPS where the sample is renewed (a) by the inclusion of a new rotation in each month and (b) by the turnover of families in dwellings which results in new households being sampled at old addresses. One hopes that the Bureau will estimate the extent of this problem, which affects both immigrants to the US population and immigrants to the household universe in the 12-month period between additions to the panel addresses.

A second problem results from following rules. Movers are followed to locations within 100 miles of the sample PSUs. This rule covers 95% of the population, but clearly selects particular types of moves. The city dwellers who tire of urban life are less well represented than persons relocating because of new employment possibilities. The Bureau needs to assess the impact of alternative following rules on the types of migration represented. CPS excludes all migration events; ISDP follows migration to the 50-mile radius; SIPP follows more extensively--These rules should be simulated on the retrospective reports of migration so that differences between the measures can be understood as differences in the population sampled and differences in response for identical populations. We need a careful evaluation.

The following rules also preclude the following of sample members into the institutional population. This rule is at odds with the principle that SIPP is a longitudinal panel of individuals living at selected addresses in Wave 1. It is understandable to question whether the instrument designed for the household population can be applied in other settings. However, some of the most important economic transitions may accompany shifts out of the household universe. The design of SIPP needs to be alert to that possibility. It must also permit measurement of persons re-entering the household universe during the life of the panel.

I heartily support Dahmann's plea for a special module of questions that is triggered by observed migration. We need to know the proximate causes of migration, the costs associated with moves, and planning for moves. No vehicle could be more appropriate for such questions than SIPP. It has a sufficient sample size so that the migrating population is large enough to study, and it can measure attitudes about migration at a time when people can easily remember.

3. Labor Force Measurements

Selection

The SIPP will be less selective of longitudinal labor force measures than matched CPS samples precisely because the CPS does not follow any movers. (In addition, CPS does not maintain adequate control over the identities of individuals so that matching is inferential, rather than positively based on an identifier such as the Social Security number.) Nonetheless selection remains a problem for SIPP. The failure to sample immigrants mentioned above, and the failure to follow household non-response from the first wave both cause systematic selection from the household population at later points in time.

The designers of the PSID express as their greatest regret that they did not attempt to follow refusals and obtain conversions. Since we do not have any data on conversion rates following a four-month lapse of time, it is hard to judge the success of such an effort. It is clear, however, that persons who refuse at any time are a portion of the population that requires diligence to sort out because they are likely to differ from cooperative respondents.

Truncation

In order to understand the implications of the complex skip structure of the labor force questions I tested the sequence on the assumption that the respondent is a student interviewed in May at the end of the school year. If he was not in the labor force during the reference period, there is no place to record student status. If the student had any work during the school term and was not underemployed, school status is not revealed. Only if additional work is desired and unavailable do we discover student status as a rationalization for not working and not looking for work. This asymmetry will make it difficult to interpret information on youth.

Another instance of truncation is revealed by the interviewer instructions. Reports of persons engaged in unpaid work are recorded only when that effort is expended in connection with the business of a family member. Work for relatives outside the household and voluntarism (for any organization) is not captured.

These problems are mentioned because they point to a set of conventions established in connection with CPS data that need not be blindly applied to an innovative survey such as SIPP.

The purpose of SIPP is to understand broadly about the economic roles of individuals and the relationship of those roles to Federal and State programs for income protection. It is widely recognized that a student role is vital to protecting the capacity of our economy to produce in the future, and that individuals investing in training are most often contributing to their own future economic well-being. It does not make sense to construe SIPP so narrowly that student roles can be understood both as complements to work roles and as an exclusive status.

Unpaid work roles are also non trivial. Persons engaged in those roles obtain skills and satisfaction and contribute to the production of the economy. We should not narrowly construct a notion of socially productive roles so that these activities are not properly recorded. SIPP could do better, without a major investment in additional measurements.

I urge the designers of SIPP to reconsider the labor force sequence in this broader setting--We want to know the contributors and the dependents of the society, not just the narrow answer to the question: "Did you receive pay for employment last week?"

Two other truncation issues should be noted. Omission of employer related data for the third employer during the reference period distorts the employment problems of a group of high turnover employees that we need to know more about. The Bureau needs to justify this arbitrary truncation of the data set. Secondly, and we have already mentioned its importance, the failure to determine the duration of labor force status at the time of the first wave of interviewing results in left-censoring of all periods of unemployment. This is most unfortunate for the study of spells of unemployment, which are of great interest to the program agencies providing unemployment compensation and other income support.

Response bias

Ryscavage stresses the fact that comparison of the 1984 and 1985 panels will expose bias due to months in sample. This is far less important than the recall biases that can be studied within the rotations of each panel. Ryscavage points to work suggesting little problem with recall bias in the annual labor market experience measures. That fact may not be relevant to the cognitive task facing individuals in the SIPP.

The year is a well-understood measure of time. Many activities relate to an annual cycle. We have many assists from nature, tax records, employment rating cycles, and other events that help us structure an answer to a question like "How many weeks did you work last year?". Most of these assists are lacking for the SIPP reference period. That fact implies that each respondent must laboriously reconstruct the last 18 weeks to answer the labor force questions. For persons with intermittent employment this may be difficult. I would argue that SIPP places a much larger cognitive burden on the respondent than the CPS reports. This will be reflected in a poorer report of events more distant in the past, and that response bias needs to be estimated and reported. I do not believe that we can interpolate extent of response errors from past observations of annual experience in relation to the CPS measure of labor force from the monthly survey.

The Four-month Report

While it is true that data are collected for a four-month reference period, it is not necessary to constrain the report from SIPP to a four-month report of labor market experience, as Ryscavage suggests. Indeed the likely recall problems with dating events four months ago, make it imperative to study reports for the most recent month. Indeed for each interview there is one report of activity during the CPS reference week of the prior month. A first task for SIPP is to validate its measures against CPS for that week.

4. Matching to Economic Data

Haber and his colleagues make it clear that a limited set of economic variables for each establishment can be matched to every worker, and a large set of variables match to those workers who are employed in large firms and particular industries. The latter data are updated annually, while data on small establishments come from periodic Censuses (selection in time) or samples (selection of firms). There appear to be five classes of potential information in Table 1.

Overlaid on this censoring of the available data is a problem of incomplete linking data in table 2.

Haber appears to be on sound ground in advocating analyses of labor supply using information on firm payrolls, size of work force, and fringe benefits. However, his suggestion that assets and capital should be imputed for small firms where only SSEL data are available is methodologically unsound. No amount of imputation will establish the covariances between worker and firm characteristics that are censored,

Table 1

Characteristic of employer: Industry	Size	Data (updating cycle, in years)
1. All	All	SSEL(1)
2. Mfg.	N½250	ASM(1)
3.	251½N½a	CM(5)
4.	a+1½N	none
5. In-scope	All	EC(5)
6.	N½500	Enterprise questionnaire

(N is the number of employees; a, an unknown parameter.)

Table 2

	Linking data in SIPP: EIN	Name	est.	Aggregation required to impute data
A.	yes	--	--	none
B.	no	yes	1	none
C.			2+	enterprise level
D.		no	--	2-digit SIC X N

by the lack of information in lines 2.-6. in Table 1 for some establishments. The only possible procedure is to study the nature of the selection to determine probability of inclusion for firms with different characteristics (Little et al. 1983).

Access to linked data

One aspect of the linked file that deserves careful thought is access by researchers outside the Census Bureau. It is clear that there is such broad scope for research using a linked file that many students of industrial organization and labor economics will wish to study the available data. Monahan (1983) has set out a model for public access using the LED. That model should be extended to all SIPP and ISDP administrative record linkages. The model for general access should include, in addition: A. On-line access by the public to a test sample (which could be transformed from real data by the addition of noise and/or the reorganization of data from real firms into simulated observations obtained by splicing vectors from several cases. B.Use of a widely available statistical package as an interface that users can manipulate to frame their questions.

5. Optimizing the SIPP Resource

SIPP is still developing as a data resource. These papers indicate that conceptual effort expended on exploiting the resource will pay off handsomely. The suggestion that SIPP be enhanced with additional questions that are conditional on change is extremely important. Not only questions that are asked following migration, but also questions that are asked following change in employers, change in family structure, and entry into the universe should be incorporated into the standard core questions. The Bureau is considering an "exit" interview to be administered to persons moving out of the sample universe; this will be extremely important both for following movers and for understanding the population on whom we have data. I should also imagine that agencies would be interested in questions that are conditional on movement on or off social welfare programs because that information would provide better clues as to the incentive effects of those programs and the processes by which individuals cope with crises in their own lives.

I strongly believe that the Bureau should consider extending the SIPP panel beyond 8 waves of interviewing. A probability sample of the individuals in each panel could be selected to be continued for an additional thirty months. If selections are made to include all the poor, persons in families undergoing demographic change, persons with unemployment, then the continuing panel would be substantially enriched with persons undergoing changes that are of interest to the social welfare and tax programs. The extension of the sample to 60 months provides a sufficient set of data that meaningful before-after studies can be made in relation to rare events such as divorce, movement off the AFDC roles, or formation of a new household.

It seems most desirable to continue discussions of the structure of SIPP and propose conditional questions, panel extension, and linkage to Census and administrative data as substantial enhancements to the current design.

References

R. J. Little, M.H. David, M. Samuhel, and R. Triest. 1983. Imputation models based on the propensity to respond, in PROCEEDINGS OF THE BUSINESS AND ECONOMICS SECTION OF THE AMERICAN STATISTICAL ASSOCIATION.

J. Monahan. 1983. Procedures for using the longitudinal establishment file. (Center for Economic Research, Bureau of the Census, Washington, D.C)

DISCUSSION

Harold Watts, Columbia University

The McMillen-Herriot paper presents a very valuable review and explorers guide for a new and relatively uncharted area. It is true that panel data have been available for some time for household samples (15 years or more) and that a literature on the topic, which is cited by the authors, goes back some 5 or 6 years. But the problem is difficult; and statistical, cross-sectional ways of thinking so dominate our approaches that a recognized solution to the longitudinal household problem has yet to be found.

Half of the two-part solution proposed here for SIPP has been implemented by the Panel Study of Income Dynamics (PSID). Both logic and experience have supported the usefulness of the "attribute" treatment of household variables. Individuals do migrate through a succession of households and treating the characteristics of the current household as an individual time-varying attribute provides a necessary perspective for analyzing persons longitudinally. But it does not solve the problem of continuity for a household. The tentative solution offered here for that task distinguishes between family and nonfamily households, imposes discontinuity whenever a unit changes from family to nonfamily status, and recognizes at most one continuation of a fissioning household. This definition seems plausible and viable, but clearly it will take some time until we have enough experience with it to make a judgement about how useful it is, or even about how useful a longitudinal household is once we fully digest how essentially ephemeral it is.

In this regard, I would like to reiterate a plea that we more carefully distinguish between "household" or "co-resident family household" and "family." The first two are ephemeral relative to individual lifetimes, at least in our society. But a family, considered as a web of kinship and socioeconomic ties and obligations, spans generations and lifetimes. I am not yet arguing that we should try to measure longitudinal family units in this dynastic sense--but lets not call something else a "family" just because we can observe it.

In terms of defining income for longitudinal household units, I am relatively confident that satisfactory solutions can be found as long as we recognize that the income picture is a lot more complicated than any "Annual Income" can begin to describe. Income is a flow concept, and we can express any measure of flow as an average annual rate no matter how long or short the period over which it has been observed. In order to aggregate properly, we have to integrate the rates with the durations, and avoid double counting, but that seems as tractible as other things we now do.

The more challenging problem is one of describing aspects of the income flow that have been totally inaccessible to the CPS--trends and instability in the rate of flow are quite important for the behavior and welfare of both individuals and households but are totally obscured in an annual income figure. Consequently, I do not feel that problems related to mimicking that very limited description of income that we have been stuck with for many years should have much impact on the way we choose to define households and describe their income experience over time.

Turning to the McNeil-Salvo paper we find a new and interesting installment of the estimated-earnings-function saga. It provides us with a new estimate of the amount of the male/female earnings differential that can be accounted for by the shorter experience and skill-depreciation that results from intervals out of the labor force. As usual, a relatively small part of the total variation in earning rates is accounted for but experience and schooling seem to have substantial and fairly stable effects for both males and females.

As measured by the authors only a small part of the male/female earnings differential can be attributed to the difference in work interruptions and consequent reduction in experience that characterizes the female work force. They estimate this by applying the male averages to female earnings function and find only about 12 percent of the gap accounted for by experience and interruption. If one goes the other way and puts the female averages into the male earnings' model, 34 percent is accounted for ($.87). I'm not sure I know which way is right--but they don't give the same answer. Still, two-thirds of the difference is left unaccounted for and only a small part of that can be attributed to the included education variables. The explanation for the difference lies in the different experience pay-off functions estimated for men and women. Especially for women with no interruption in work, the pay-off for experience is extremely small.

I would like to comment more generally on the potential use of SIPP for this kind of study. This particular study does not take advantage of the longitudinal features in the ISDP, and I feel strongly that future work of this kind should exploit the power of these data more fully.

One of the hypotheses related to the interruption of work phenomenon has to do with rate of recovery of "depreciated" earning rates. ISDP and SIPP both provide several readings, over time, on earnings rates, and hypotheses about rates of change can be confronted directly. A longitudinal data base with historical recall questions can also do much more than indicate current marital status. Duration and cumulative effect of statuses like marriage and recency of change can be brought to bear on more highly articulated models for testing implications of one or another theories about the earnings/experience relationship.

It is unfortunate for these particular issues that the SIPP measure of the length of interruptions remains somewhat crude even in the latest questionnaire. An interruption "from" 1962 "to" 1962 can be anything from 6 to 12 months in duration; from 1962 to 1963, the range is 6 to 24 months; from 1962 to 1964, it is 14 to 36 months. Such imprecision dilutes whatever effect this variable may have. It would seem that a question aimed at getting a start date and duration or an end-point and duration would be equally feasible and provide a more definite interval of time.

SURVEY OF INCOME

AND

PROGRAM PARTICIPATION:

SESSION III

This section is comprised of five papers presented in this session which was sponsored by the Section on Survey Research Methods.

OBTAINING CROSS-SECTIONAL ESTIMATES FROM A LONGITUDINAL SURVEY: EXPERIENCES OF THE INCOME SURVEY DEVELOPMENT PROGRAM

Hertz Huang, Bureau of the Census

I. INTRODUCTION

In 1975 the Secretary of the Department of Health, Education and Welfare (The Department of Health and Human Services (HHS) predecessor agency) authorized a program, the Income Survey Development Program (ISDP), to resolve technical and operational issues for a major new survey --the Survey of Income and Program Participation (SIPP). Much of the work of the ISDP centered around four experimental field tests that were conducted in collaboration with the Bureau of the Census to examine different concepts, procedures, questionnaires, recall periods, etc. Two of the tests were restricted to a small number of geographic sites; the other two were nationwide. Of the two nationwide tests, the more important data collection was called the 1979 Research Panel. This panel consisted of nationally representative samples which provided a vehicle for feasibility tests and controlled experiments of alternative design features. Information concerning the ISDP may be found in Ycas and Lininger (1981), David (1983), and the survey documentation now available through the National Technical Information Service (1983).

The 1979 Research Panel was a multiple frame sample consisting of a general population (area) sample of 9300 initially designated addresses drawn from the 1976 Survey of Income and Education (SIE) and some Census Bureau's current survey reserve measures and new construction update, and two list frame samples of (a) eligible applicants for the Basic Educational Opportunity Grant (BEOG) Program and (b) blind and disabled Supplemental Security Income (SSI) recipients.

The 1979 Research Panel was a longitudinal survey consisting of six waves of interviewing. The sample was divided into three interviewing panels. The first panel was first interviewed in February 1979; the second panel was first interviewed in March; and the third panel was first interviewed in April. Each sample unit was subsequently interviewed every three months. A sample of addresses was chosen and persons living in the sample units (addresses) during the first wave of interviews were defined as original sample persons. For interviews subsequent to the first, the sample of addresses became a sample of persons; accordingly, original sample people were followed to their new addresses in subsequent interviews (with reasonable geographic constraints -- within 50 miles of any ISDP Primary Sampling Unit). Personal interviews were conducted in Wave 1 with all adults (persons sixteen years old and over) at the sampled address. These become the original sample persons. During Waves 2-6 all persons currently residing with an original sample person were interviewed. This means, for example, that if an original sample person moved to a new address with four other adults, then questionnaires were administered to everyone at the original sample person's new address. If any original sample person remained at the first wave address, anyone who moved into that address with the original sample person was also interviewed. Thus, interviews were conducted with all adults at an address as long as at least one of the adults present was an original sample person. Because of the ISDP rules, persons can be lost from sample because they move beyond the survey's boundaries; in addition, people were added to the sample because they became part of the housing unit in which the original sample person resides.

Obviously, the universe changes continuously through the life of the survey. A great deal of interest exists, however, in developing cross-sectional estimates at the time of each interview wave. In the absence of drawing a new sample at each interview, any cross-sectional estimates developed for Waves 2-6 are subject to a population coverage bias. This paper will focus only on the covered population and present some unbiased base weights for cross-sectional estimators for the non-institutionalized U.S. population represented by the longitudinal sample (the population coverage bias will remain, however, since units containing no persons who were in the universe at the time of Wave 1 cannot come into sample). Since the methodology for treating both area sample and list frame samples was needed for ISDP 1979 Research Panel, both will be described below. The estimation methods described here are directly applicable to the Survey of Income and Program Participation (SIPP), an overall description of which is found in Nelson, McMillen, and Kasprzyk (1984) and Herriot and Kasprzyk (1984).

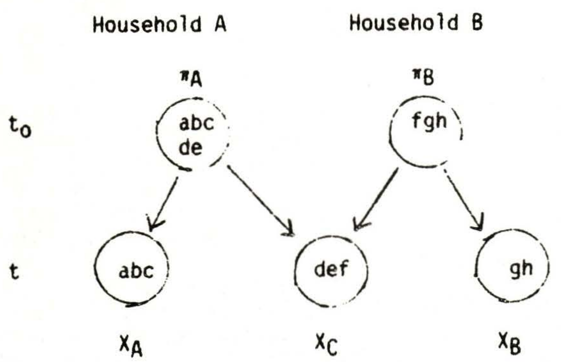
II. THE POPULATION FOR CROSS-SECTIONAL ESTIMATES

We begin by defining the general population for which estimates are required. All households existing during the first wave of interviews (February through April 1979) are considered the initial population. Based on the rules adopted for the following individuals who move, we have essentially a longitudinal sample of persons as well as households for the initial population. Since no new sample was drawn at any subsequent interview, the sample does not completely represent the non-institutionalized U.S. population after first quarter of interview. There were persons in the following categories at the initial interview time but became part of the non-institutional population at a subsequent wave of interviewing: 1) U.S. citizens living abroad, 2) citizens of other countries who subsequently move to the U.S., 3) persons in institutions or armed forces barracks. These persons will be called the group R subpopulation which did not have chance to be selected as original sample persons. At a subsequent wave of interviews, the longitudinal sample did not include any household in which all current members were in the group R subpopulation. However, persons in the group R subpopulation who later joined households that included original persons eligible for sampling in the first wave were added to the cross-sectional universe. these persons along with newborns will be called "additions" in subsequent waves. In general, "additions" are defined as persons moving into eligible households after the first wave who were not eligible for sampling in the first wave.

III. GENERAL CONCEPT OF CROSS-SECTIONAL ESTIMATION

Due to the procedures adopted for following movers in the 1979 Research Panel, at subsequent interviews a household could consist of members from more than one household in the universe at the time of the first wave. The inclusion probability of such a household would depend on the inclusion probabilities of the households which the members of the current household were part of at the time of the first interview. The inverse of the inclusion probability is usually used as the weight of a sample household in estimation. However, because of the sample design of the 1979 Research Panel, the inclusion probability of a household is a function of its primary sampling unit, type of sampling frames and the 1975 income of the household which occupied the housing unit during the SIE interview. Only the inclusion probability of an original sample household was feasible to calculate. The inclusion probability of an original nonsample household is almost impossible to evaluate, but such households can come into sample on later waves. Therefore, some alternative weighting procedures needed to be explored.

The idea to be presented in this discussion is very simple. We will associate observations at any given point in time with the known inclusion probabilities of the original sample households. We will split up observations belonging to a household when current household members come from more than one original household. A portion of the observation is then associated with each original household. The following example will illustrate the idea: Assume that A & B are two original households with inclusion probabilities π_A and π_B respectively. At the first wave of interviews, household A consists of five members, a,b,c,d, and e, and the household B consists of three members, f,g, and h. During the second wave of interviews we find that d,e, and f are living together and form a new household, called household C, while a,b, and c are still in household A and g and h are still in household B.



Two alternatives are proposed, both involving the division of household C into two parts; one part is associated with household A and the other with household B.

a) Multiplicity Approach:

Based on the number of ways (called multiplicity) that the new household C can be included in the sample, the observation (additive, such as counts, income or values) of household C (called X_C) is divided by the number of original households involved (two in this case) and each portion is added to the corresponding observation of household A (called X_A) and household B (called X_B). Therefore, if both households A and B are original sample households, the cross-sectional estimate, $\hat{X}$, for the total at the second wave based on these three households can be expressed as:

$$\hat{X} = \frac{1}{\pi_A}\left(X_A + \frac{1}{2}X_C\right) + \frac{1}{\pi_B}\left(X_B + \frac{1}{2}X_C\right)$$

$$= \frac{1}{\pi_A}X_A + \frac{1}{\pi_B}X_B + \left(\frac{1}{2\pi_A} + \frac{1}{2\pi_B}\right)X_C.$$

Hence, the weight for the new household c is $\frac{1}{2\pi_A} + \frac{1}{2\pi_B}$. If only household A is an original sample household, then the weight for the new household is $\frac{1}{2\pi_A}$; if only household B is an original sample household then the weight for the new household is $\frac{1}{2\pi_B}$.

b) Fair Share Approach

This approach assumes that all household members contribute equally to their household. Thus, the observation of household C is divided into appropriate portions based on the proportion of members of household C which come from each original household (2/3 from household A and 1/3 from household B in this example). Therefore, if both households A and B are original sample households, the cross-sectional estimate for the total at the second wave based on these three households is expressed as

$$\hat{X} = \frac{1}{\pi_A}\left(X_A + \frac{2}{3}X_C\right) + \frac{1}{\pi_B}\left(X_B + \frac{1}{3}X_C\right)$$

$$= \frac{1}{\pi_A}X_A + \frac{1}{\pi_B}X_B + \left(\frac{2}{3\pi_A} + \frac{1}{3\pi_B}\right)X_C.$$

Hence, the weight for the new household c is $\frac{2}{3\pi_A} + \frac{1}{3\pi_B}$. If only household A is an original sample household, then the weight for the new household is $\frac{2}{3\pi_A}$; if only household B is an original sample household then the weight for the new household is $\frac{1}{3\pi_B}$.

Since our sample was longitudinal in nature, if the universe remained constant through time, original sample persons would be a representative sample of the universe at any given point in time. Hence, using the inclusion probabilities of the original sample persons, the above estimators are unbiased (proof is given in next section). However, our feasible target population (excluding the group R subpopulation) changes through time by including the 'additions' (defined in Section II). To compensate for this, we will include these "additions" in the proposed estimators below.

IV. PROPOSED ESTIMATORS FOR GENERAL POPULATION (AREA) FRAME

Before the estimators are given, some notation should be defined. Note that some of the defined quantities may not be observed. For the first wave of interview (time t_o), let

$X(t_o) = \sum_{k=1}^{N(t_o)} X_k(t_o)$ the parameter to be estimated, where $X_k(t_o)$ is the value of the characteristic for the k^{th} unit (which may be a person or a household) and $N(t_o)$ is the number of units in the universe at time t_o;

α_k = 1 if unit k was in the sample at time t_o, k = 1, ..., $N(t_o)$
= 0 otherwise

π_k = the probability that unit k was selected for the sample at the first wave of interview (time t_o)
= $P_r(\alpha_k = 1) = E(\alpha_k)$, k=1,..., $N(t_o)$ At a subsequent wave (time t), define for each household i:

S_i = the total number of current residents of household i at time t

r_i = the number of original eligible households from which the current residents come (does not include households from which "additions" come)

and

$S_{i1}, S_{i2}, \ldots, S_{ir_i}$ be the number of current residents in household i from each of the r_i original households and S_{ia} be the number of current residents from the "additions" as defined in Section II. Write

$$S_i = \sum_{j=1}^{r_i} S_{ij} + S_{ia} \quad \text{and} \quad S_{io} = \sum_{j=1}^{r_i} S_{ij}$$

Also define N(t) as the total number of units in the target population at time t, such as household units (include all the original households plus newly formed households) or other units based on a group of persons such as families or sub-families (include all sample persons interviewed nonsample persons plus "additions"). And let $X(t) = \sum_{i=1}^{N(t)} X_i(t)$ be the parameter (total) to be estimated for the target population at time t. To simplify the notation, we will replace N(t), X(t) and $X_i(t)$ by N, X and X_i respectively.

Based on the general concept described in Section III, two cross-sectional estimators are proposed for the area frame to estimate the total of a characteristic of the target population at time t.

Estimator I (Multiplicity Estimator):

This estimator is based on the multiplicity of each current household

$$\hat{X}_M = \sum_{i=1}^{N} W_{Mi} X_i$$

where

$$W_{Mi} = \sum_{j=1}^{r_i} \frac{\alpha_j}{r_i \pi_j}.$$

Note that α_j and π_j are associated with original households but are reindexed within each current household i. It is easily seen that

$$E(\hat{X}_M) = E(\sum_{i=1}^{N} W_{Mi} X_i) = E\left(\sum_{i=1}^{N} \sum_{j=1}^{r_i} \frac{\alpha_j}{r_i \pi_j} X_i\right)$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{r_i} \frac{E(\alpha_j)}{r_i \pi_j} X_i = \sum_{i=1}^{N} \frac{1}{r_i} \left(\sum_{j=1}^{r_i} \frac{\pi_j}{\pi_j}\right) X_i = \sum_{i=1}^{N} X_i = X$$

Therefore $\hat{X}_M$ is an unbiased estimator of X. Note also that if $\alpha_j = 0$ it is not necessary to know Π_j, so that W_{Mi} can be calculated based on the selection probability only for sample units.

Estimator II (Fair Share Estimator):

This estimator is motivated by the assumption that all current household members contribute equally to the household in which they reside for the major household characteristic values, such as earnings and welfare benefits.

$$\hat{X}_F = \sum_{i=1}^{N} W_{Fi} X_i$$

where

$$W_{Fi} = \sum_{j=1}^{r_i} \frac{S_{ij}}{S_{io}} \frac{\alpha_j}{\pi_j}$$

As in the multiplicity estimator, α_j and π_j are associated with original households but are reindexed within each current household i. One can see that $\hat{X}_F$ is also an unbiased estimator of X as follows:

$$E(\hat{X}_F) = E\left(\sum_{i=1}^{N} W_{Fi} X_i\right)$$

$$= \sum_{i=1}^{N} \frac{1}{S_{io}} \left(\sum_{j=1}^{r_i} \frac{S_{ij} E(\alpha_j)}{\pi_j}\right) X$$

$$= \sum_{i=1}^{N} X_i = X$$

Note that if household j was not in sample at time t_o, it is unnecessary to know the number of current residents from original household j, S_{ij}, in $\hat{X}_F$ since $\alpha_j = 0$. Also note that because "additions" are not included in the weight calculations, they must be identified and excluded from using either estimator.

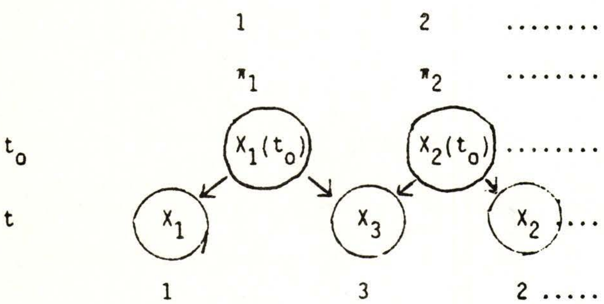
Comparison of Estimator I and Estimator II

Both Estimator I, $\hat{X}_M$, and Estimator II, $\hat{X}_F$, are feasible to compute. We now compare them with respect to both operational convenience and reliability.

In order to compute $\hat{X}_M$, the number of original households eligible for sampling from which the current residents come is needed. This information is particularly difficult to obtain at each successive wave of the survey. However, to compute $\hat{X}_F$ one only needs to know the number of current residents in the household (excluding new additions) and the number of residents from each original sample household. This information could be obtained from the 1979 Research Panel person identifier without collecting additional information.

The equal contribution from the members of a household is a natural assumption. It reflects better the actual share among the household members in the absence of knowledge of the actual contribution from each member. For example, without knowledge of each person's age, employment status and other needed information, it is more logical to assume that earnings and welfare benefits are equally contributed by household members than any arbitrary way of defining household members' shares. And as will be seen below, that heuristically $\hat{X}_F$ can be justified as the approximate minimum variance unbiased estimator under what seems to be natural assumptions given a state of ignorance about the actual shares of the household members.

Assume that at a subsequent wave time t three households are generated from two original households of the first wave of interview (time t_o) as follows:



Let $X_k(t_o)$, k=1, ..., $N(t_o)$ be the value of the characteristic of interest for household k at time t_o and X_j, j=1, ..., N be that value for household j at time t. Using Section III we divide up X_3 in two parts, fX_3 and $(1-f)X_3$ and then associate fX_3 with X_1 and $(1-f)\,X_3$ with X_2. Without loss of generality, assume that a sample size of 1 was selected at the first wave, t_o, with probability π_k, k=1, ..., $N(t_o)$. An unbiased estimator of total, X, at time t can be written as

$$\hat{X} = \frac{\alpha_1}{\pi_1} X_1 + \frac{\alpha_2}{\pi_2} X_2 + \left(\frac{f\alpha_1}{\pi_1} + \frac{(1-f)\alpha_2}{\pi_2} \right) X_3 + \ldots$$

where α_i, i=1, $N(t_o)$ is defined at the beginning of this section. Notice that both $\hat{X}_M$ and $\hat{X}_F$ are special cases of $\hat{X}$. The variance of $\hat{X}$ is

$$\mathrm{Var}(\hat{X}) = \pi_1 \left\{\left(\frac{1}{\pi_1} X_1 + \frac{f}{\pi_1} X_3\right) - X\right\}^2$$

$$+ \pi_2 \left\{\left(\frac{1}{\pi_2} X_2 + \frac{1-f}{\pi_2} X_3\right) - X\right\}^2 + \ldots\ldots$$

The remaining terms are not explicitly given here since they are not functions of f. The $\mathrm{Var}(\hat{X})$ is minimized if

$$f = \left(\frac{X_2+X_3}{\pi_2} - \frac{X_1}{\pi_1} \right) \Big/ \left(\frac{X_3}{\pi_1} + \frac{X_3}{\pi_2} \right)$$

Since usually not both π_1 and π_2 are known and in most of the surveys conducted by the Bureau of the Census, the inclusion probabilities, π_k, are about the same for all ultimate sampling units (even though they are unequal in the 1979 ISDP), one may simplify f to $f = (X_2 + X_3 - X_1)/2X_3$

Obviously, a weight defined as a function of survey observations is not easy to implement. To further simplify f, we assume the percentage growth of X from t_o to t is constant for all units involved and define

$$a\, X_1(t_o) = X_1 + X_{31}$$

$$a\, X_2(t_o) = X_2 + X_{32}$$

$$X_3 = X_{31} + X_{32}$$

where X_{3i} is the share of X_3 belonging to household members from original household i, i=1,2.

Without knowledge of both $X_1(t_o)$ and $X_2(t_o)$, one might naturally assume that the two initial households are about the same i.e., $X_1(t_o) = X_2(t_o)$ and reduce f to $X_{31}/(X_{31} + X_{32})$.

Now if the contribution to X_3 is proportional to the number of persons from each original household, then $f = S_{31}/S_{30}$, as defined in W_{Fi}. This result can be extended to any sample size as well as to the case that the new household members are from more than two original households. Therefore, without knowledge of the actual contribution from each household member, $\mathrm{Var}(\hat{X}_F)$ is smaller than $\mathrm{Var}(\hat{X}_M)$ under these assumptions.

V. PROPOSED ESTIMATORS FOR LIST FRAMES

Since persons are the list frame sampling units, we can divide all persons in the general population into three groups based on their relationship with the list frame under consideration.

I) Persons who are included in the list frame (called list frame persons). For the SSI list frame, this group includes all the (under 65) recipients of the Federal Supplemental Security Income in December 1978; while for the BEOG list frame, this group includes all the eligible applicants of the Basic Educational Opportunity Grant as of September 1978 for school year 1978-79.

II) Persons who are not included in the list frame but live with a list frame person(s) during the first wave of interview (February through April 1979).

III) Persons who are not included in the list frame nor do they live with a list frame person(s) during the first wave of interview.

Both Group I and II had some chance to be included in the list frame sample, but Group III did not. The original (first quarter) households which consist of Group I and/or Group II persons will be called list frame households. As time went on, some members of Group III moved in and lived with person(s) belonging to Group I or II. Such members of Group III will be 'additions' for the list frame, since they are not initially eligible for sampling in the list frame. Note that the type of persons already described as "additions" for the general population (as defined in Section II) will also be "additions" for the list frame. For the following discussions, we now define two types of "additions" for the list frames: the "additions" that come from Group III will be called "Group III additions" and the type of "additions" as defined for the area frame will be called "area frame addition."

If a list of recipients of a government assistance program is used as a list frame then Group III is usually fairly large. If we construct our estimators the same way as we did for the area frame, we will include many of Group III persons in our estimates at time t of subsequent interviews. Consequently, we wouldn't really know what "subpopulation" we were estimating. In our opinion, it is not feasible to define such a subpopulation at time t. Without new sample drawn each wave from the updated list, a proper cross-sectional estimate for a list frame subpopulation at time t is not likely, especially if the turnover rate of the list frame members is high. Therefore, we will restrict our cross-sectional estimates to be based on only the original list frame sample persons (that is, the list frame persons selected for list frame sample plus all the persons who reside with them during the first quarter of interview) and the "area frame additions." In so doing we know that at any time t, the target population we are estimating consists of the original list frame subpopulation (that is Groups I and II) and the type of "additions" as defined in the area frame. Note that the original list frame subpopulation is determined by persons who were on the list at the time of sample selection. They may not be on the list by the time of initial interview.

In the 1979 ISDP panel, a household may have a multiple chance of being selected for the list frame sample if more than one member of the list frame persons live in that household at the first wave of interview. (Some effort was made to reduce multiple chance of selection for those households in SSI frame.) Therefore, the concept of the base weight for the first wave of interview is no longer trivial.

Similar to the area frame, we define $X(t_0) = \sum_{i=1}^{N^L(t_0)} X_i(t_0)$ as the population parameter to be estimated from a list frame sample at time t_0, where $X_i(t_0)$ is the value of the characteristic for the i^{th} unit in the list frame subpopulation, which includes both Group I and II defined at the beginning of this section. Let

α_i = 1 if list frame person i is in the sample,
= 0 otherwise (note that $\alpha_i = 0$ for all non-list frame persons)

π_i = the probability that list frame person i is selected for the list frame sample for the first wave of interview (time t_0)
= $Pr(\alpha_i = 1) = E(\alpha_i)$

β_{oj} = the number of list frame persons (indexed by i) in the j^{th} household at time t_0.

Then the base weight at time t_0 for the j^{th} household and its residents is defined as

$$W_{oj} = \sum_{i=1}^{\beta_{oj}} \frac{\alpha_i}{\beta_{oj}\pi_i}$$

where α_i and π_i are associated with list frame persons but are reindexed within household j.
For time t of a subsequent wave, let

β_k = the total number of list frame persons living in the original (time t_0) list frame households which the current residents of the k^{th} household come from.

S_k = the total number of current residents at time t; $S_{k1}, S_{k2}, \ldots, S_{kr_k}$ be the number of current residents in the k^{th} household who come from each of r_k original list frame households; S_{ka} is the number of current residents of the k^{th} household who are from the "area frame additions"; and $S_{k\,III}$ is the number of current residents of the k^{th} household who are from the "Group III additions." Therefore

$$S_k = \sum_{j=1}^{r_k} S_{kj} + S_{k\,III} + S_{ka}$$

$$= S_{kc} + S_{ka}.$$

N^L = the total number of units such as household or family units, in the list frame universe at time t

The two cross-sectional estimators for the total of a characteristic of the list frame target population at time t are as follows:

Estimator I (Multiplicity Estimator)

To avoid estimating "Group III additions" we will treat all the current residents from the "Group III additions" as a separate list frame sampling unit. Therefore, in the k^{th} household at time t, there are $\beta_k + U_k$ list frame sampling units, where $U_k = 1$, if some of the current residents in the k^{th} household are from "Group III additions," 0 otherwise. The multiplicity estimator for the list frame population total is given in the following:

$$\dot{X}_M^L = \sum_{k=1}^{N^L} W_{Mk}^L X_k$$

where

$$W_{Mk}^L = \sum_{i=1}^{\beta_k} \frac{\alpha_i}{(\beta_k + U_k)\pi_i}.$$

α_i and π_i are associated with original list frame person but are reindexed within each current household k.

Estimator II (Fair Share Estimator)

Motivated by the assumption that all current residents contribute equally to a household we propose the following list frame estimator:

$$\dot{X}_F^L = \sum_{k=1}^{N^L} W_{Fk}^L X_k$$

where

$$W_{Fk}^L = \sum_{j=1}^{r_k} \frac{S_{kj}}{S_{kc}} W_{oj};$$

and S_{kj} and W_{oj} are associated with original household but are reindexed within each current household k.

These two estimators are constructed to estimate the list frame subpopulation excluding the "Group III addition." They are not unbiased estimators in the global sense, independent of the value of the characterisic of interest. However, the fair share estimator would be unbiased under the assumptions that all current residents contribute equally to a household and a household is treated as a fraction of a household if some of the current residents are from "Group III additions."

In addition to the "unbiasedness" described above, $\hat{x}_F^L$ is also preferred for the same reasons (operational and reliability) stated in the area frame. In computing $\hat{x}_F^L$, we need to know β_{0j}, the number of list frame persons in a sample household at the initial interview (time t_0). This information was not difficult to obtain. And at any subsequent wave of interview time t, we needed to know only S_{kc}, the total number of current residents who are not "area frame additions" and S_{kj}, the number of current residents from each original list frame sample household. The latter can be obtained through the person identifier.

However, in order to compute $\hat{x}_M^L$ at time t we would have to ask a complicated question to obtain β_k, the total number of list frame persons living in the original households from which the current residents come.

VI. SUMMARY

These two estimators were constructed based on the specific procedure of following movers in the 1979 ISDP. However, they can be easily modified to apply to other designs and procedures. The fair share estimator was actually used for the 1979 ISDP. It is also being used for the 1984 Survey of Income and Program Participation.

As noted in Section III, the inverse of the inclusion probability of a household at time t is usually used as the weight of that household to obtain an unbiased estimator. When a household consists of members from two original households (called households i and j), the inclusion probability of this new household is $\pi_i + \pi_j - \pi_{ij}$, where π_{ij} is the joint selection probability of households i and j at the first wave of interview. This inclusion probability is operationally impossible to obtain, but its inverse can be reduced to the weight (W_{Mi}) of the multiplicity estimator in most surveys conducted by the Census Bureau. In these surveys, the wave 1 inclusion probabilities are almost the same for all ultimate sampling units and the joint selection probabilities are generally due to the sampling without replacement within PSUs which are generally negligible. Therefore, the fair share estimator not only overcomes the trouble of obtaining such inclusion probabilities, but it has good variance properties under some reasonable conditions and it is easy to implement.

As described in Section V, the application of this approach to multiple frame longitudinal surveys presents additional problems, and the resulting estimators are not nearly as satisfactory.

This research was completed before the first interviews of the 1979 ISDP Research Panel. Horvitz and Folsom (1980) have done similar work in conjunction with the National Medical Care Utilization and Expenditure Survey.

VII. ACKNOWLEDGMENTS

The aurthor gratefully acknowledges the many helpful comments of Dr. D. Kasprzyk and other colleaques at the Census Bureau.

REFERENCES

[1] David, M. (ed), (1983). Technical, Conceptual and Administrative Lessons of the Income Survey Development Program. New York: Social Science Research Council.

[2] Herriot, R. and Kasprzyk, D. (1984). The Survey of Income and Program Participation. To be published in the Proceedings of The Social Statistics Section, American Statistical Association.

[3] Horvitz, D.G. and Folsom, R.E. (1980). Methodological Issues in Medical Care Expenditure Surveys. The Proceedings of the Section on Survey Research Methods, American Statistical Association.

[4] Nelson, D., McMillen, D. and Kasprzyk, D. (June 1984). An Overview of the Survey of Income and Program Participation. SIPP Working Paper Series, Number 84-01.

[5] Survey Documentation for 1979 Research Panel, (1983). Department of Commerce, National Technical Information Service, Springfield Virginia.

[6] U.S. Department of Commerce, Bureau of the Census, (1978). Some initial thoughts on Weighting the 1979 Income Survey Development Program. Memorandum from G. Capps to M. Schooley, November 30, 1978.

[7] U.S. Department of Commerce, Bureau of the Census, (1979). Some Alternative Procedures on Weighting the 1979 Income Survey Development Program. Memorandum from M. Schooley to G. Shapiro, March 1, 1979.

[8] Whitmore, R., (ed.) (1982). Family Unit Weighting Methodology for the State Medicaid Household Survey Components of the National Medical Care Utilization and Expenditure Survey. Report Prepared for National Center for Health statistics and Health Care Financing Administration.

[9] Ycas, M. and Lininger, C., (1981). The Income Survey Development Program: Design Features and Initial Findings. Social Security Bulletin 44:11, November 1981.

WEIGHTING OF PERSONS FOR SIPP LONGITUDINAL TABULATIONS

David Judkins, David Hubble, Rajendra Singh, David McMillen, Larry Ernst
Bureau of the Census

I. INTRODUCTION

Since October of 1983, the Census Bureau has been conducting interviews for a new survey, the Survey of Income and Program Participation (SIPP). The survey will effect long-sought improvements in the measurement of annual income and the complex relationships between income flows, labor force participation, participation in government programs such as welfare, and tax policy. One of the products of the interviewing will be a set of longitudinal records on a probability sample of the population. The subject we address in this paper is the weighting of these longitudinal records so that the data may be analyzed.

We are aware of only two precedents for this weighting. They are the National Medical Care Expenditure Survey (NMCES) and the National Medical Care Utilization and Expenditure Survey (NMCUES). The latter was conducted jointly by the Research Triangle Institute and the National Opinion Research Center[2]. Some work was done on the problem for the Income Survey Development Program (ISDP)[6], but it was not implemented. The techniques used by them are among those under consideration for SIPP. Naturally though, we are also considering some new ideas. These ideas are still in a very preliminary form. We are presenting them here to get early reaction and suggestions from the statistical community.

Our general approach consists of three major steps. The first step is to derive an unbiased weight for each longitudinal record. This is not as straightforward as it seems due to the fact that a slightly different set of people is being interviewed each month. Section III discusses this step.

The second step is to make adjustments for those records that are incomplete. We will use imputation when part of an interview is missing. (See Samuhel's paper in this session [3].) We will also probably use imputation when a whole interview is missing where the missing interview is bracketed by good interviews. Our research on adjusting for records with more than one missing interview is in too preliminary a stage to report on. (One proposal has been made by Little and David[4].)

The third step is to correct for disproportional representation of demographic types to reduce variance and gain some consistency with the Current Population Survey (CPS). Section IV discusses this step.

Before discussing the weighting, it is essential that we define which of the many possible longitudinal universes is the universe for which estimates are to be provided. Section II deals with this problem.

Finally, we mention some of the important features of the design of SIPP. For more details, the reader is encouraged to first read an overview of the survey [5]. Roughly 20,000 households were interviewed between October 1983 and January 1984, inclusively. That set of interviews constitutes the first wave of the 1984 panel of SIPP. The Census Bureau will try to interview the persons in those households an additional seven or eight times in four-month waves, even if they move. We will also interview any persons who "usually reside" with anyone in the original cross-section for at least one-half of a calendar month. This extra interviewing will only be conducted for the time period that the joint residence is maintained. Only the original cross-section is followed through moves.

II. DEFINING THE LONGITUDINAL UNIVERSE OF PERSONS FOR SIPP

The SIPP universe at the beginning of any panel is persons who are members of the civilian non-institutional population, and members of the military not living in barracks on bases. Defining the longitudinal universe is somewhat more complicated. We begin by defining the possible ways persons can enter and exit this universe. Next we discuss the relationship between the cross-sectional universes and the longitudinal universe. The third topic of this section addresses the definition of table universes, and a discussion of calculating annual income for persons in the longitudinal universe.

There are two ways persons can enter the SIPP universe: 1) persons can move from overseas (immigrate or return), institutions, or from military barracks; 2) persons can be born to members of the universe.

Similarly, there are two methods of exiting the universe; 1) moving overseas, to an institution, or to military barracks 2) dying. Given these conditions of entering and exiting the universe, and a definition of the initial universe, we can define the universe at any subsequent point in time, and the means by which the universe grows and diminishes over time. The next problem is to make the transition from the cross-sectional universes to a single longitudinal universe.

There are three methods of defining a longitudinal universe: 1) the composition can be fixed at some point in time; 2) the universe may be defined as the union of some set of cross-sectional universes; and 3) the universe may be defined as the intersection of some set of cross-sectional universes.

A longitudinal universe may be defined at a given point in time. For example, we can take the civilian noninstitutional population at the time the sample is drawn, at the midpoint of the panel duration, or at the end of the panel to define the universe of interest. Of course, the time point chosen could be any time point within the duration of the panel. This rather narrow definition of the universe has an advantage in its simplicity, but also several disadvantages. Dependent on the chosen point in time, this definition produces a strictly declining population, a first increasing and then decreasing population, or a strictly increasing population. In the first case all entrants are excluded from the longitudinal universe, and only exits are allowed to alter the universe. In the second case, entry is allowed and exit is denied until the midpoint, when the situation reverses. In the last case, all those who exit during the panel are excluded from the longitudinal universe and only entries are allowed to alter the universe. In addition, it is difficult to argue why one point or another should be chosen as the point in time to define the universe, and for some purposes you may need a different point than the one originally chosen.

The next two definitions build from the above idea that a universe may be defined at any point during the panel. Let us assume then a set of universes each defined at a different point in time. To further simplify discussion, let us assume a set of twelve monthly universes defined at the midpoint of each month. The two options are to use either the union or the intersection of these sets.

Consider first the union of sets. The union of these monthly universes is all persons who were at some point during the year members of the civilian noninstitutional population. In other words, all members of the target population plus all persons who enter or exit during the year are included in the union of sets definition. This is the most inclusive of the universe definitions offered here, and the one which best captures the dynamic characteristics of the population. Some of the disadvantages of this type of definition will be raised below in the discussion of tabulations and table universes.

An alternative to the union of sets is the intersection of the set of twelve monthly cross-sectional universes. Here we include in the longitudinal universe only those persons who were members of all of the cross-sectional universes. In other words, only those persons who were members of the civilian noninstitutional population or the special military categories on the fifteenth of each of the twelve months. This definition is even more restricted than the point-in-time definition. This intersection of sets definition produces a static population. That is to say there is no entering or exiting allowed.

Of the three longitudinal definitions offered here, only the union of sets incorporates the dynamic qualities that are inherent in a longitudinal process.

That would seem to make it the logical choice; however, this is also the definition that produces the most complications when tabulating data. Consider, for example, tabulating marital status at the beginning of the year with marital status at the end of the year. There is no place in such a table for persons who were in universe at one point in time, and not in the universe at the other point in time. For the union of sets definition there is a need for both a column and a row for persons not in the universe at time 1 or not in universe at time 2. For those definitions that allow exiting only a column for persons not in the universe at time 2 is necessary as long as the beginning point of the universe and the tables are the same.

Similar problems arise in computing annual income. Aggregating across months is simple, but it is not clear how to compare income amounts for full year and part year persons. That is simply to say that a $6,000 income for 6 months and a $6,000 income for 12 months are not the same.

III. INITIAL WEIGHTING

For SIPP, as for ISDP, a cross-section of the population will be followed for a period of time. Data will also be collected on the people that the original cross-section live with. The original idea was that only the data on the people in the original cross-section would be used in person longitudinal tabulations; the data on the other people would be used only to provide the "household experience" of the original cross-section. We are now reexamining that idea. The data on the other people can be used to better understand the experience of new entrants to the SIPP universe. Furthermore, there are ways to use these data more intensively to gain valuable variance reductions. Unfortunately, these procedures require strong assumptions for unbiasedness. In the following sections, we explore the trade-off. We first discuss whether the data on the other people should be used. We then discuss how to construct weighting procedures that use these data more or less intensively.

A. Variance Reduction Versus Bias Control.

Let us first define some terms and clarify the type of parameters to be estimated. We divide the sample people into three groups. A person is an original sample person if he/she is a member of the original cross-section.[1] A person is an associated sample person if he/she was a member of the eligible population at the time the cross-section was selected but happened not to be selected. Anyone else is an additional sample person. This last group consists of recent discharges from institutions, new immigrants, and people moving out of military barracks. The type of parameter to be estimated is the frequency of some pattern of labor force participation, program participation, income receipt, etcetera, by demographic characteristics, housing characteristic, geographical unit, educational background, etcetera. A simple example is the frequency of women who were receiving public assistance in January 1984 but were not

---

[1] A person in original cross-section of households who was 15 years old or older at the time of the first interview is definitely an original sample person. Twelve, thirteen, and fourteen year old children are more difficult to classify. At first, no questionnaires are filled out for them and they are not followed in the rare event of an unaccompanied move. However, after they turn 15, they are treated the same as any other original sample person. We will treat them here as original sample people. Children eleven or younger are not classified at all.

receiving it in December 1984.

The original idea was to estimate parameters like this one by summing the weights of all original sample persons with the desired characteristics. Data on associated and additional sample people are needed only to classify original sample people with respect to household characteristics; for example, was the original sample person living in a household in which at least one member received social security? Given this scheme, no data are needed on associated or additional sample people for the period that they don't reside with original sample people. Hence, we do not follow associated or additional sample people if they separate from original sample people. Clearly then, the data on associated and additional sample people are frequently incomplete.

Despite this incompleteness, we are now considering ways to squeeze more information out of this data. The first way is to provide estimates for the "union" universe using the data on additional sample. The second way is to use the data on both types to reduce variances. To begin the argument for this second use, we first point out that for shorter time periods these data are frequently either complete or nonexistent. (Throughout this section, by complete we mean complete ignoring nonresponse.) This is always true for 1 month periods, usually true for 3 month periods, and frequently true for 12 month periods. For example, suppose that Ruth is an original sample person interviewed in October 1983. In November, she marries Jack, who was in the October SIPP universe. They stay together at least through April 1985. Then Jack is an associated sample person on whom we have complete 1984 data. Alternatively, suppose that Jack was living in a military barracks in October 1983. Then he is an additional sample person on whom we have complete 1984 data. There will obviously be many more cases in these complete categories for 1985 data. Furthermore, there will be many cases where we are only missing one or two months of data.

Intuitively, it seems wasteful to give zero weights to these cases with complete or almost complete data, as originally intended. On the other hand, zero weights must be assigned to the seriously incomplete cases to avoid large-scale imputation. One possible solution is obtained by initially assigning strictly positive weights to all cases, including these that are incomplete due to field procedures, and then treating the incomplete cases as if they were caused by non-response. Imputation would be used for the almost complete cases. Note then that the seriously incomplete cases would have zero weights, while the other cases would have positive weights. If enough data has been collected on the associated and additional sample people to correctly model the probability of this type of nonresponse, then we would still have unbiased estimators.

An example of the type of model required is that starting from a given social-economic stratum, the new economic situation of a male divorcee does not depend on whether he or his ex-spouse was the original sample person. Here we stress that if a person has responded to even a single wave of SIPP, then we have an extraordinary wealth of data available for modeling.

Future Study

Of course, we will never know for certain whether such a model is correct. There is a risk of biasing the estimators, and as a rule the Bureau is willing to risk biases for decreases in variance only if there is some evidence that the bias squared is substantially less than the variance decrease. Our plans at this time are not well formulated. A reasonable first step is to quantify for each proposed weighting procedure the frequency of positively weighted incomplete cases by the severity of the incompleteness. The only source for this information is the ISDP. We are currently working on ways to get appropriate tabulations for it.

B. Construction of Unbiased Weighting Procedures

Below we present a very simple result that characterizes a general class of unbiased procedures. Reflection on this result quickly helps one to understand that there are infinitely many unbiased procedures. Most of them are totally inappropriate, but it is very possible that better and radically different weighting procedures exist than have yet been conceived.

Let $x = \sum_{1}^{N} x_i$ be the parameter of interest to be estimated where x_i is the value of the characteristic for the i^{th} unit. Let w_i be a random variable associated with the i^{th} unit such that $E(w_i) = 1$.

$$E(Y) = E(\sum_{1}^{N} w_i x_i) = \sum_{1}^{N} E(w_i) x_i = \sum_{1}^{N} x_i = x.$$

If the probability of selection is known for all units, it is common to take

$$w_i = \begin{cases} \text{inverse probability of selection if } i^{th} \text{ unit is in sample;} \\ 0 \text{ otherwise.} \end{cases}$$

This definition of w_i is not, however, necessary. In this case it is impossible since the probabilities are unknown.

Each sample person has a cross-sectional weight for every month that they are in the universe. These cross-sectional weights have expected value of unity, are strictly positive for the months that the person is in sample, and are zero for the months that interval but after the mid date are assigned their respective cross-sectional weights as of the date they enter it, as in Procedure 1 and 2. Persons who leave the universe before the "mid" date are assigned their respective cross-sectional weights as of the date they leave it.

Procedure 4. Average Cross-Sectional Weight (ACS)

Each person recieves a longitudinal weight valid for a specific time interval. Persons that remain in the universe throughout the interval are assigned the average of their respective monthly cross-sectional weights. Persons that enter or leave the universe are assigned the average of their respective monthly cross-sectional weights for the months they were in the universe during the time interval. Positive weights are assigned to all sample persons. A more formal definition is given below.

Let U_i = number of months the i^{th} person was in the universe during the specified time interval

Let C_i = sum of the monthly cross-sectional weights of the i^{th} person in the specified time interval

Then the person longitudinal weight is C_i/U_i.

C. Comparison of Procedures

In this section we describe in detail the types of complete and incomplete cases that are used by each procedure. First, we need to define some notation. Let

t_B = the first month that a person is in the universe,

t_E = the last month that a person is in the universe,

t_1 = the first month that a person is in sample,

t_2 = the last month that a person is in sample,

t_m = the mid-month of the interval of interest.

The description is given in Table 1. The first 14 cases comprise the "intersection" universe. The remaining 32 cases fill out the "union" universe. Each case is marked as having complete, partial or no data for the interval of interest. Of course, all of this is assuming perfect response. The only type of missingness that we are discussing here is that caused by operational procedures. On the right, there is a column for each procedure with an "X" if the procedure uses the case.

The entry date procedure uses the perfect cases 1,15,17, and 18, but does not use the perfect cases 2 and 16; the partial cases 3,5, and 19-27; and cases 12 and 44 for which no relevant data exists. The beginning date of interval and mid date of interval procedures both use all of the perfect cases, more of the partial cases and none of the completely missing cases. We thus think that these two procedures will tend to yield smaller variances than the entry date procedure with possibly some small increase in the risk of bias. The average cross-sectional procedure is the the person is not in sample. By choosing the longitudinal weight to be the cross-sectional weight at a particular time or the average of the cross-sectional weights at several points in time, we can construct longitudinal weighting procedures that use different subsets for the overall data set.

In this section we present four longitudinal weighting procedures for computing unbiased estimates for persons. They are all presented in terms of the "union" universe, but they can be easily modified for the "intersection" universe by assigning a zero weight to any person who is not in every one of the 12 cross-sectional universes. In Section III.C we compare the procedures with respect to the use of data collected on associated sample persons and additional sample persons. In the full paper there is an additional section with examples of the application of these procedures.

Procedure 1. Entry Date Weight (ED)

Each person receives a single longitudinal weight for any time interval that contains at least part of the period for which the person was in the universe, namely the cross-sectional weight for the person at his/her entry date into the universe. For all original and associated sample persons, the entry date into the universe is the start of the panel, so their longitudinal weights are their Wave 1 cross-sectional weights. For those who enter the universe after Wave 1, (additional sample persons), the longitudinal weight is the cross-sectional weight of the household, of which they are a member, as of the date they enter the universe. If the cross-sectional weight of the household at that date is zero, then the additional sample person's longitudinal weight is zero.

Procedure 2. Beginning Date of Time Interval Weight (BDI)

Each person receives a longitudinal weight valid for all time intervals with the same beginning date. Persons in the universe at the beginning date of the time interval are assigned their respective cross-sectional weights for that date. Persons that enter the universe during the time interval are assigned their respective cross-sectional weights as of the date they enter it, as in Procedure 1.

Procedure 3. "Mid" Date of the Time Interval Weight (MDI)

This procedure is similar to Procedure 2. Each person receives a longitudinal weight valid for a specific time interval. Persons in the universe at the "mid" date of the time interval are assigned their respective cross-sectional weights at that date. The difference is that instead of the person longitudinal weights being determined at the beginning date of the time interval, these weights are determined at some predesignated date within the time interval. Persons that enter the universe during the time most aggressive in utilizing partial data. It uses all the perfect and partial cases and none of the completely missing cases. Also note that it assigns smaller weights, in general, to the partial cases than the perfect cases. We think it will tend to yield the smallest variances with the greatest risk of bias.

IV. CONTROLS

We are currently considering the adjustment of SIPP longitudinal weights so as to achieve the variance reductions associated with ratio estimation while also causing agreement with SIPP cross-sectional controls on a monthly basis; i.e., in addition to simple undercoverage adjustments we are considering the possibility of forcing the sum of the longitudinal weights of all persons in the universe in a given month to equal the cross-sectional population control for that month. Since longitudinal weights are fixed over time while the universe fluctuates over time, such agreement will not occur unless proper steps are taken to ensure it. We are also considering adjustments to force spouses to have equal longitudinal weights. We are considering these two possibilities in order to enhance the face validity of the survey at the least possible cost of reduced precision.

Objectives

The primary reason for ratio adjustment of longitudinal weights is to reduce variances of longitudinal weights by ensuring representativeness with respect to demographic variables which are highly correlated with the variables to be measured. (This is frequently referred to as post-stratification.) To the extent that it corrects for differential undercoverage, it is also hoped that bias is reduced by ratio adjustment.

A reasonably good adjustment is to proportionately adjust the weights of persons by demographic type in a specified month so that the weighted counts agree with independent population estimates by demographic type for that month. Persons not in sample in the chosen month are assigned the factor for their demographic type. This approach operates under the assumption that the degree to which the sample represents each demographic type is not highly variable over time. This adjustment does not adjust weights to monthly controls other than those for the chosen month. Another approach is to make the adjustment for all persons for each of the 12 data months, then assign to a person the average of the 12 factors for his/her cell. Such an adjustment would tend to be influenced less by the vagaries of sample selection.

Addressed here is the more complex problem of adjusting weights for disproportional representation in a manner such that consistency with cross-sectional controls is achieved for each month. This problem has a multitude of solutions. However, the solution we seek should be the one which provides the greatest variance reduction. One possible solution is to first adjust weights as outlined in the above paragraph, then further adjust them so that the desired monthly consistency is achieved while minimizing the amount by which weights are further adjusted. This can be done with Lagrange multipliers or with linear programming. This approach preserves the benefits of the initial adjustment by demographic variables provided that this second adjustment causes relatively small changes in weights. Research is needed to determine whether the second adjustment would indeed cause only small changes.

A further refinement would be to adjust so that spouses have equal weights. Naturally, persons undergo changes in marital status during the year; some persons may have more than one spouse over a one year period. Define a "marriage group" to be a group of persons in the SIPP sample, each of whom has been or is married to at least one other person in the group during the data year. It is possible to perform an adjustment so that all persons in a given marriage group have equal weights. This last adjustment would cause slight disagreements between longitudinal population estimates and monthly controls; it appears likely that such disagreements could be made arbitrarily small by iteratively repeating the two adjustment steps for consistency with cross-sectional estimates and consistency within marriage groups. For more details, see our full paper.

REFERENCES

[1] Census Bureau memorandum from C. Jones to T. Walsh, "Cross-Sectional Weighting Specifications for the First Wave of the 1984 Panel of the Survey of Income and Program Participation (SIPP)," November 25, 1983.

[2] Jones, Bruce L., "Development of Sample Weights for the National Household Survey Component of the National Medical Care Utilization and Medicare Utilization and Expenditure Survey," April 1982.

[3] Samuhel, Michael E., "Longitudinal Item Imputation in a Complex Survey," presented to the Survey Research Methods Section of the American Statistical Association during the 1984 Annual Meetings.

[4] Little, R.J.A. and David, M., "Weighting Adjustments for Nonresponse in Panel Surveys," 1983 Working Paper.

[5] Nelson, D., McMillen, D., and Kasprzyk, D., "An Overview of the Survey of Income and Program Participation," SIPP Working Paper Series No. 8401. U.S. Bureau of the Census, Washington, D.C. 1984.

[6] Kasprzyk, D. and Kalton, G., "Longitudinal Weighting in the Income Survey Development Program," in <u>Technical, Conceptual and Administrative Lessons of the Income Survey Development Program (ISDP), Papers presented at a conference, October 6-7, 1982</u>. Social Science Research Council, Washington D.C., 1983.

[7] Sirken, Monroe G., "Household Surveys with Multiplicity," <u>Journal of the American Statistical Association</u>, 65, No. 329 (1970), 257-66.

Table 1. Case Utilization by Procedure

Case	Preceeding Time Interval	Interval of Interest	Succeeding Time Interval	Completeness	Procedure ED	BDI	MDI	ACS
1	$t_B=t_1$		$t_2 \le t_E$	Perfect	X	X	X	X
2	$t_B<t_1$		$t_2 \le t_E$	"		X	X	X
3	$t_B=t_1$	$t_m \le t_2$	t_E	Partial	X	X	X	X
4	$t_B<t_1$	$t_m \le t_2$	t_E	"		X	X	X
5	$t_B=t_1$	$t_2<t_m$	t_E	"	X	X		X
6	$t_B<t_1$	$t_2<t_m$	t_E	"		X		X
7	t_B	$t_1 \le t_m$	$t_2 \le t_E$	"			X	X
8	t_B	$t_1 \le t_m \le t_2$	t_E	"			X	X
9	t_B	$t_1 \le t_2 < t_m$	t_E	"				X
10	t_B	$t_m<t_1 \le t_2$	t_E	"				X
11	t_B	$t_m<t_1$	$t_2 \le t_E$	"				X
12	$t_B=t_1 \le t_2$		t_E	No Data	X			
13	$t_B<t_1 \le t_2$		t_E	"				
14	t_B		$t_1 \le t_2 \le t_E$	"				
15	$t_B=t_1$	$t_2=t_E$		Perfect	X	X	X	X
16	$t_B<t_1$	$t_2=t_E$		"		X	X	X
17		$t_B=t_1$	$t_2 \le t_E$	"	X	X	X	X
18		$t_B=t_1$ and $t_2=t_E$		"	X	X	X	X
19	$t_B=t_1$	$t_m \le t_2 < t_E$		Partial	X	X	X	X
20		$t_B=t_1$	$t_2<t_E$	"	X	X	X	X
21		$t_B=t_1$ and $t_m \le t_2$	t_E	"	X	X	X	X
22		$t_B=t_1$ and $t_m \le t_2 < t_E$		"	X	X	X	X
23	$t_B=t_1$	$t_2<t_m \le t_E$		"	X	X		X
24	$t_B=t_1$	$t_2<t_E<t_m$		"	X	X		X
25		$t_B=t_1 \le t_2 < t_m$	t_E	"	X	X		X
26		$t_B=t_1 \le t_2 < t_m \le t_E$		"	X	X		X
27		$t_B=t_1 \le t_2 < t_E \le t_m$		"	X	X		X
28	$t_B<t_1$	$t_m \le t_2 < t_E$		"		X	X	X
29	$t_B<t_1$	$t_2<t_m \le t_E$		"		X		X
30	$t_B<t_1$	$t_2<t_E \le t_m$		"		X		X

Case	Preceeding Time Interval	Interval of Interest	Succeeding Time Interval	Completeness	Procedure ED	BDI	MDI	ACS
31	t_B	$t_1 \le t_m \le t_2 < t_E$		Partial			X	X
32	t_B	$t_1 \le t_m$ and $t_2 = t_E$		"			X	X
33		$t_B < t_1 \le t_m$	$t_2 \le t_E$	"			X	X
34		$t_B < t_1 \le t_m \le t_2$	t_E	"			X	X
35		$t_B < t_1 \le t_m \le t_2 \le t_E$		"			X	X
36	t_B	$t_1 \le t_2 < t_m \le t_E$		"				X
37	t_B	$t_1 \le t_2 < t_E \le t_m$		"				X
38	t_B	$t_m < t_1 \le t_2 \le t_E$		"				X
39		$t_B < t_1 \le t_2 < t_m$	t_E	"				X
40		$t_m < t_B < t_1$	$t_2 \le t_E$	"				X
41		$t_m < t_B < t_1 \le t_2$	t_E	"				X
42		$t_B < t_1 \le t_2 < t_E \le t_m$		"				X
43		$t_m \le t_B < t_1 \le t_2 \le t_E$		"				X
44	$t_B = t_1 \le t_2$	t_E		No Data	X			
45	$t_B < t_1 \le t_2$	t_E		"				
46		t_B	$t_1 \le t_2 \le t_E$	"				

LONGITUDINAL FAMILY AND HOUSEHOLD ESTIMATION IN SIPP

Lawrence R. Ernst, David L. Hubble, and David R. Judkins, Bureau of the Census

1. INTRODUCTION

Many types of statistics will be produced by the Survey of Income and Program Participation (SIPP), but there is one type that was the driving force behind the unique design of the survey. To be fully successful, SIPP must tell us what happens to households over the course of time. From it we must obtain estimates of the patterns of income receipt, program participation, and labor force participation at the household and family level by a host of other characteristics. Of particular interest are parameters such as total annual household income and the number of families that have stopped drawing food stamps by demographic characteristics.

Before estimates can be produced, a decision must be made on the definition of a longitudinal household to be used in this survey. (To simplify the presentation, we will concentrate our discussion on longitudinal households as opposed to longitudinal families. However, parallel longitudinal estimation procedures can readily be developed for families). It often happens that the occupants of several housing units move and regroup. We need to know which, if any, of the resulting households are to be considered continuations of the previous households. Many definitions have been proposed, but final agreement has thus far not been achieved. Also decisions have yet to be made on whether households that form or dissolve during a time interval of interest are to be considered as part of the universe for estimation purposes. Because of the absence of agreement in these areas, several proposed definition and universe combinations will be considered in this paper. They are listed in Section 2. Also because of this absence of agreement, the major aim of this paper will be simply to compare several possible longitudinal household estimation procedures and present criteria for choosing among them, without attempting to reach a conclusion on a preferred procedure.

We foresee several steps in the process of producing longitudinal households estimates. The focus in this paper is the first step, the production of weights that would yield unbiased estimates assuming there are no data that are missing or in error, and that the frame coverage is perfect. Several procedures for obtaining such weights will be presented in Section 3. Choosing among these procedures is complicated by the fact that even assuming perfect response, data needed to produce unbiased estimates will be missing for some households because they are not collected with the current field procedures. This difficulty is prinipally due to the fact that, except for a few household definitions, all unbiased procedures assign positive weights to some longitudinal households for time periods when they are not in sample. The severity of this problem and the extent to which it is correctable in the future by changing field procedures or by modeling the missing data, vary by procedure. This problem, along with descriptions of other important features, both positive and negative, that estimation procedures may possess is presented in Section 4. Finally, in Section 5 a detailed comparison of the features of the estimation procedures under consideration in this paper is presented.

It is assumed in this paper the reader has a basic knowledge of SIPP, including the design of this survey. Nelson, McMillen, and Kasprzyk (1984) provides this information.

Portions of the original paper, principally an examples section and a section on adjustments to the unbiased weights, are omitted here due to lack of space. The complete paper is available from the authors.

2. LONGITUDINAL HOUSEHOLD DEFINITIONS

In this section three possible longitudinal household definitions are presented to illustrate the longitudinal weighting procedures that will be described in the next section. (A fourth definition, known as the Shared Experience Definition was included in the original paper, but omitted here due to lack of space. In terms of the properties discussed in Section 5 it is identical to the Reciprocal Majority Definition that is included here.) A thorough discussion of longitudinal household definitions is presented in McMillen and Herriot (1984). In addition, several other terms will be defined, including the longitudinal household universes considered in this paper.

Since household composition and data for SIPP are obtained on a monthly basis, each of the definitions to be presented will be in terms of household continuity from one month to the following month. A longitudinal household over a time interval of n (≥ 2) months is then defined to be one which is continuous for each of the n-1 corresponding pairs of consecutive months. (It has not yet been decided if this approach will actually be used in SIPP.)

For each of the definitions below the conditions for which household B at month t+1 is the continuation of household A at month t are stated. One condition that we require that all the definitions share is that A and B are either both family households or both nonfamily households. The other conditions are:

No Change Definition (NC). A and B have the same household members.

Same Householder Definition (SH). A and B have the same householder. As an alternative, householder could be replaced by principal person in this definition without altering any of the statements made about it in subsequent sections, provided the final estimation procedure in Section 3 is also modified accordingly. (The householder of a household is, roughly, the person who owns or rents the housing unit. The principal person is the wife in a married-couple household, and the householder in all other households.)

Reciprocal Majority Definition (RM). The majority of individuals who are both household members of A at time t and in the universe at time t+1 are members of B at time t+1, and the majority of individuals who are both household members of B at time t+1 and in the universe at time t are members of A at time t. (This type of longitudinal definition was originally developed by Dicker and Casady (1982) for use in the National Medical Care Utilization and Expenditure Survey (NMCUES).)

We will now clarify several other terms.

A household is said to be in existence over a time interval of $n \geq 2$ months if it is longitudinal over that time interval. Its period of existence is the longest such time interval. In the case of a household which is defined cross-sectionally for a month t, but is not longitudinal over either of the two month intervals containing t, then the period of existence of the household is defined to be one month.

If t_1 and t_2 are any pair of months, and longitudinal estimates are to be made over the interval $[t_1, t_2]$, then the following two possibilities will be considered in subsequent sections for the universe of households for which estimates will be produced.

Restricted Universe(R). The set of all households in existence over the entire interval $[t_1, t_2]$.

Unrestricted Universe(U). The set of all household in existence for one or more months in $[t_1, t_2]$.

Each sample panel is interviewed eight times. Each of the eight rounds of interviews takes four consecutive months to complete and is known as a wave.

Finally, we define an original sample person to be a person that was in sample during the first wave and will be at least 15 years of age by the end of the panel.

3. UNBIASED WEIGHTING PROCEDURES

In this section we present five weighting procedures for computing estimates of totals or proportions for longitudinal households that would be unbiased in the sense that the expected value of the estimator over all possible samples is the parameter of interest assuming no data are missing or in error, and perfect frame coverage. Modifications and adjustments of these estimation procedures necessary because of the unrealistic nature of these assumptions are considered in the original paper, but are omitted here due to lack of space. Except for the Continuous Household Members procedure, which will only be applied to the restricted universe, all the procedures will be stated for the unrestricted universe. To apply them to the restricted universe simply zero weight each household which is not in continuous existence over the time interval of interest. Furthermore, unless otherwise stated, all the procedures will be applied to all four longitudinal definitions defined in Section 2.

First we will explain why a common method of estimation, weighting by the reciprocal of the probability of selection is not feasible for our purposes, and hence the need to consider alternative procedures. Let $X = \sum_{i=1}^{N} x_i$ be a parameter of interest, where x_i is the value of the characteristic for i-th unit in a population of size N. Typically in survey work, to estimate X a sample would be drawn in such a manner that the i-th unit has a known positive probability p_i of being chosen, and X would then be estimated by

$$\hat{X} = \sum_{i=1}^{N} w_i x_i , \tag{3.1}$$

where

$$w_i = \begin{cases} \dfrac{1}{p_i} & \text{if the i-th unit is in sample,} \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

Unfortunately for household and family estimation in SIPP, both cross-sectionally and longitudinally, such an estimation approach is not practical. For example, cross-sectionally a household is interviewed and used in the estimation process for a given month if and only if at least one household member is an original sample person. Consequently, to use (3.1) and (3.2) as an estimator it would be necessary to determine the probability that at least one member of the current household is an original sample person. It would be operationally impossible to determine this probability, since it would first be necessary to determine the first wave households for all current household members and then compute the probability that at least one of these first wave households was selected.

Fortunately though, it is not necessary that w_i satisfy (3.2) in order that (3.1) be unbiased. In fact if w_i is any random variable associated with the i-th unit in the population satisfying

$$E(w_i) = 1, \tag{3.3}$$

then (3.1) is unbiased, that is $E(\hat{X}) = X$. Thus, defining unbiased longitudinal household and family weighting procedures reduces to defining random variables w_i satisfying (3.3).

Before we present the longitudinal weighting procedures we will state what, for purposes of this paper, a cross-sectional household weight is, since most of longitudinal weighting procedures will be defined in terms of cross-sectional weights. The first wave cross-sectional weight for a sample household is taken here to be the reciprocal of the probability of selection. For all nonsample households in the universe this weight is defined to be zero. For any month after the first wave a different definition is necessary because of possible changes in household composition. So, the cross-sectional household weight for any such month is defined to be the mean of the first wave cross-sectional household weights for all persons in the household that month who will be at least 15 years of age by the end of the panel and who were in the universe during the first wave. This type of weighting procedure is currently being used in SIPP to produce cross-sectional estimates, hence the name. It is readily verifiable that the weights satisfy (3.3).

We also will leave it to the reader to verify that the weights for each of the longitudinal procedures to be presented satisfy (3.3) and hence lead to unbiased estimators.

Beginning Date of Household Procedure (BH). Each longitudinal household receives a single weight valid for any time interval that contains at least part of the period for which the household existed, namely the cross-sectional weight for the household at the beginning date of the household. In particular, if there were no original sample persons in a household at its beginning date then its longitudinal weight would be zero. This approach to longitudinal household estimation was previously used in the NMCUES (Whitmore, Cox and Folsom 1982).

Beginning Date of Time Interval Procedure (BI). Each longitudinal household receives a longitudinal weight valid for all time intervals with the same beginning date, namely the cross-sectional weight for the household at the beginning date of the time interval. Longitudinal households that form during the time interval are assigned the cross-sectional weight for the household at its beginning date, as in the preceeding procedure.

Continuous Household Members Procedure (CM). The following procedure will only be applied to the restricted universe, as defined in Section 2. For any time interval for which the household is in existence the longitudinal weight to be assigned is determined by the set of persons that are members of the household throughout the time interval. The longitudinal household weight is the cross-sectional weight that would be assigned to a household consisting of this set of persons; that is, the average of the first wave weights of these people. A longitudinal weight of zero is assigned to the household if there are no original sample persons who are members throughout the time interval. The procedure is slightly biased because a longitudinal household with no members continuously present throughout a time interval has no chance of receiving a positive weight, thereby making satisfaction of (3.3) impossible. Since we believe this situation will rarely occur, at least for the longitudinal household definitions considered here, we expect this bias to be very small.

Average Cross-Sectional Household Weight Procedure (AW). Each longitudinal household receives a longitudinal weight valid for a specific time interval, namely the average of the monthly cross-sectional weights for the household over the intersection of the life of the household and the specified time interval.

Note, there are many procedures, like AW, that entail the averaging of weights, both household cross-sectional weights and person longitudinal weights. We will examine only one of these procedures here, as an example of this type of longitudinal household weighting procedure.

Householder Weight Procedure (HW). The following procedure will be applied only to the No Change and Same Householder Definitions, since it is appropriate only for definitions that allow for a single householder during the household's existence. (Generalizations of this procedure which are not so restricted in their applicability exist but will not be considered here.) The procedure assigns a single weight valid for any time interval that contains at least part of the period for which the household existed, namely the first wave cross-sectional household weight of the householder's first wave household. A longitudinal weight of zero is assigned to the household if the householder was not an original sample person.

As will be seen in Section 5, this procedure is clearly the one of choice when the Same Householder Definition is used. If that type of definition is used with householder replaced by principal person then a similar modification of this estimation procedure with householder replaced by principal person would be appropriate.

4. POTENTIAL ADVANTAGES AND DISADVANTAGES

The ideal unbiased weighting procedure would provide a single set of weights applicable to any time interval, require no more data than were collected, and possess the minimum variance among all unbiased procedures. Unfortunately, no such procedure exists. The procedures described in Section 3 all fail one or more of these three criteria to various degrees. In this section, we explain the nature of the failures without explicitly comparing the procedures. That is done in Section 5.

Multiplicity of Weights. Some procedures have the advantage of assigning to each household a single weight which depends only on conditions as of the first reference month for the household and which is valid for every interval that the household is in the universe. Other procedures have the disadvantage of sometimes producing different weights for the same household for different time intervals. (Procedures with this disadvantage could be modified so that only a single weight applies to any time interval, by computing for each household the weight appropriate for that procedure for the unrestricted universe and the 2 1/2 year time interval corresponding to the life of the panel. The weight obtained would also be used for any smaller subinterval for which the household is in the universe. However, weights obtained in this manner might not be able to be determined until the end of the life of the panel. This would make them difficult to use because we would have to wait until the last data from the panel were processed before estimates could be produced for any earlier time period. In any case, such weights would often lead to higher variances for short time intervals than weights developed specifically for the short time intervals.)

Unavailable Data Requirements. Most definition and procedure combinations require data from some households for time periods when the household is in existence but not in sample, that is for time periods for which interviews are not conducted for the household because no original sample people are members of the household. This needed data could be information for determining proper longitudinal weights or subject-matter information for use in tabulating the estimates. Some of this information is not collected for the 1984 panel of SIPP because of the current operational procedures. This is a consequence of the fact that agreement has not been reached on the longitudinal household definition to be used in SIPP. In this vacuum, operational procedures were determined mainly by considerations of difficulty and cost. Once a definition has been agreed on, depending on the nature of the unavailable data, it might be possible to change operational procedures for future SIPP panels so that the required data are collected. To understand the problem with current operational procedures, consider the following situation. A household is longitudinal from month t_B to t_E. Original sample people are part of the longitudinal household only from month t_1 to t_2. If $t_B<t_1$, then some prior information may be unavailable. Revised operational procedures to obtain this information might involve retrospective questions, longer reference periods or proxy data on anyone who left the household before the first interview. If $t_2<t_E$, then some posterior information may be unavailable. Revised operational procedures might involve interviewing the household through t_E.

One of the important discriminants between the weighting procedures is how successfully they avoid the need for data from the period that the longitudinal household exists but is not in sample. (The need for such data is avoided by assigning zero weights to these problem households.) In terms of information needed for weighting, some procedures require only enough data to determine whether $t_B<t_1$, while others need to know t_B even when it is less than t_1. Similarly, some procedures only require knowledge of whether $t_2<t_E$, while others need to know t_E even when it is greater than t_2. Furthermore, besides this need for information for determination of weights, if any parameters other than the number of longitudinal households are to be estimated, then required subject-matter data may be missing as well, either before t_1, after t_2, or both.

While the problem of missing information is a serious one, it is not fatal. Procedures can be developed to compensate for the unavailable data. Specifically, the data collected on these households while they were in sample should be sufficient for performing imputation for existence/nonexistence outside the in-sample period and formation and/or dissolution dates. The imputed values can then be used to calculate weights for these households. These households can then be treated as noninterviews so that the weights of mover households with similar demographic characteristics but with complete data receive increased weights while the deficient households themselves receives zero weights.

If the models underlying the procedures developed for adjusting for the missing information are true then it is still possible to obtain unbiased estimators, although now in a model-based sense. Furthermore, since the missing information that we are concerned with here is not caused by refusal to respond, modeling in this context might not suffer from the usually imperfect assumptions on similarity between respondents and nonrespondents that underlie any adjustments that use data from respondents to account for data missing from refusals. In addition, because of the longitudinal nature of the survey, there is generally a large amount of data available from the problem households that could be used in such adjustments. However, if the models are not perfect, then in general, the larger the proportion of data required that is unavailable, the greater the potential for serious bias problems.

Variances. In general, estimation procedures with the smallest variances are those that utilize available data intensively and tailor the weights to the specific time interval of interest. Unfortunately, as shall be seen in the next section, such procedures are often characterized by heavy needs for unavailable data which, as noted above, may impact unfavorably upon bias. Thus, there often is a direct trade-off between variance and the risk of bias. It will be difficult to weigh these factors against each other, since it appears that no single procedure will provide the correct balance for all of the multitude of characteristics that will be estimated by SIPP.

For use in the next section, we will define some labels for the advantages and disadvantages identified in the foregoing discussion. Let:

- T_1 mean that a single longitudinal weight exists for each household, valid for all time intervals for which the household is in the universe, and which depends only on conditions which could be determined during the first interview,
- T_2 mean the negation of T_1,
- BW_1 mean that no data from the period preceeding the first interview are unavailable but required for weighting,
- BW_2 mean that we need to know for weighting whether the longitudinal household existed before the first interview,
- BW_3 mean that we need to know for weighting the conception date of the household (within the time interval of interest),
- BD_1 mean that no subject-matter data from the period preceeding the first interview are unavailable but required,
- BD_2 mean the negation of BD_1,
- FW_1 mean that no data from the period following the last interview are unavailable but required for weighting,
- FW_2 mean that we need to know for weighting the dissolution date of the household (within the time interval of interest),
- FD_1 mean that no subject-matter data from the period following the last interview are unavailable but required,
- FD_2 mean the negation of FD_1.

Note that T_1, BW_1, BD_1, FW_1 and FD_1 are the desirable properties.

5. DETAILED COMPARISONS OF ADVANTAGES AND DISADVANTAGES

Table 1 below presents advantages and disadvantages of each definition, procedure and universe combination. A comparison of these features follows the table. Next, an explanation of each entry in the table is given. Finally, a discussion of data utilization, which is not in Table 1, is presented.

Table 1.

Definition	Procedures	Universe	T_1	T_2	BW_1	BW_2	BW_3	BD_1	BD_2	FW_1	FW_2	FD_1	FD_2
			Features										
NC	All	Both	X		X			X		X		X	
SH	HW	Both	X		X			X		X		X	
SH, RM	BH	U	X			X		X		X			X
SH, RM	BH	R	X			X		X			X		X
SH, RM	BI	U		X		X		X		X			X
SH, RM	BI	R		X	X			X			X		X
SH, RM	CM	R		X	X			X		X		X	
SH, RM	AW	Both		X			X		X		X		X

Comparison of Features in Table 1. As noted at the end of Section 4, T_1, BW_1, BD_1, FW_1, and FD_1 are the desirable properties. For the NC definition all five procedures considered here possess all these desirable properties, as does the HW procedure for the SH definition.

However, for the SH and RM definitions, and most other definitions too, the BH, BI, and CM procedures have different subsets of the set of desirable features, so that the procedure to be adopted depends, at least in part on the features deemed most important. AW possesses none of these desirable features for these two definitions. Its principal advantage lies in possible reductions in variances because of complete utilization of available data, which will be discussed later. BH has advantages T_1, BD_1, and FW_1 for the unrestricted universe, and T_1 and BD_1 for the restricted universe. The main reason for consideration of this procedure would be that it is the only one among BH, BI and CM that always has advantage T_1. BI has advantages BD_1 and FW_1 for the unrestricted universe and BW_1 and BD_1 for the restricted universe. Its principal advantage over BH is that for the restricted universe no retrospective questions need be asked. CM (which is only applicable to the restricted universe) possesses all desirable features except T_1, that is no information not currently collected is needed for this procedure. Recall, however, that CM had the disadvantage of being slightly biased as explained in Section 3.

Explanation of Entries in Table 1. All explanations presented below apply to both universes unless otherwise stated.

NC Definition, All Procedures. Since the composition of a household is unchanged throughout its period of existence under NC, we have the following two possibilities:

(a) No original sample people were in the household at any time during its period of existence, in which case the longitudinal household weight is zero for any time interval and procedure.

(b) One or more original sample people were in the household throughout its existence, in which case the beginning and ending dates of the household are known, as is the composition of the household and complete data for each month of its existence. Consequently, features BW_1, BD_1, FW_1, and FD_1 apply.

Furthermore, T_1 applies since procedures BH, BI, CM, and AW all reduce to the cross-sectional household weight at the beginning date of the household, while HW is the weight of the householder at the beginning date.

SH Definition, HW Procedure. The explanation is similar to the one given above, except now the two cases are: (a) The householder was not an original sample person. (b) The householder was an original sample person.

SH and RM Definitions, BH Procedure. T_1 is applicable, since by definition the weight is the cross-sectional household weight as of the beginning date of the household. BW_2 applies because the longitudinal household weight is the cross-sectional household weight as of the first month in sample if the household began that month, while otherwise the weight will be zero since there were no original sample people in the household when it began. (For the restricted universe, households which entered sample after the beginning of the time interval always receive a zero weight.)

BD_1 holds since all households with positive weights were in sample at their beginning date and no retrospective subject-matter data is therefore needed.

FW_1 holds for the unrestricted universe since the weight is determined at the beginning date of the household. However, for the restricted universe, it is necessary to know if the household continued to exist throughout the entire time interval because it receives a zero weight for the time interval if it did not continue. Under current procedures a household which no longer has any original sample person is not followed, and it would therefore generally not be possible to determine if it remained in existence for the entire time interval. Consequently, FW_2 applies.

FD_2 applies since there would be missing data for all households with positive weights which continued to exist after there were no longer any original sample people present, which could happen for either of these definitions.

SH and RM Definitions, BI Procedure. T_2 is applicable since time intervals with different beginning dates may yield different longitudinal weights. BW_1 applies for the restricted universe, since the longitudinal weight is the cross-sectional household weight as of the first month of the time interval for all households in sample that month, and zero for all other households. However, BW_2 applies for the unrestricted universe since longitudinal households that entered sample after the beginning of the time interval are treated as in the BH procedure.

BD_1 holds since any household with a positive weight was either in sample the first month of the time interval or the month that the household began, and consequently, no retrospective data are needed.

As in the BH procedure, and for the same reasons, FW_1 applies for the unrestricted universe, FW_2 for the restricted universe and FD_2 for both universes.

SH and RM Definitions, CM Procedure, Restricted Universe. T_2 is applicable since any two intervals may yield different longitudinal weights.

Furthermore, BW_1, BD_1, FW_1, and FD_1 apply. The explanation is similar to that given for the NC definition except now the two cases are: (a) No original sample people were household members for the entire time interval. (b) At least one original sample person was a household member for the entire time interval.

SH and RM Definitions, AW Procedure. For an explanation for this row of the table see the original paper.

Utilization of Data. Having compared the procedures with respect to needs for unavailable data and the multiplicity of weights, we now turn our attention to variance. To compare the variance characteristics of the procedures we will focus on the amount of collected data that is used in obtaining estimates, since this is a primary determinant of variance. This discussion will also better illustrate the proportion of data needed for estimation that is unavailable for each procedure. In general, the greater this proportion is, the larger the burden is on any missing data procedure employed, with a resulting greater potential for bias problems. To make the comparison we show in Table 2, all 24 possible cases of how the data on a longitudinal household may be complete, partly available, or nonexistent for a particular time interval.

The symbols t_B, t_1, t_2, and t_E denote beginning date of household, first sample month, last sample month, and ending date of household respectively. The columns indicate different time intervals. Interval B is the interval of interest. Interval A is from t_B until the beginning of interval B, while interval C is from the end of interval B until t_E. The fifth case, for example, is of a household that formed before interval B about which we are missing some data pertinent to the early part of interval B. The first nine cases comprise the restricted universe. The last 15 cases fill out the unrestricted universe. Each case is marked as having complete data, partial data, or no data. Of course, all of this is assuming perfect response. The only type of missingness that we are discussing here is that caused by operational procedures. On the right there is a column for each procedure with an "A" entered if it always uses the case, an "S" if it sometimes uses the case but not always (which will be explained in the discussion that follows), and a blank otherwise. These comparisons do not apply to the NC definition, for which all five procedures use all the complete cases and no other cases.

Table 2.
Data Utilization

	Interval A	Interval B	Interval C	Completeness	Procedure BH	BI	CM	AW	HW
1	$t_B=t_1$		$t_2<t_E$	perfect	A	A	S	A	S
2	$t_B<t_1$		$t_2<t_E$	perfect		A	S	A	
3	$t_B=t_1$	t_2	t_E	some missing	A	A		A	
4	$t_B<t_1$	t_2	t_E	some missing		A		A	
5	t_B	t_1	$t_2<t_E$	some missing				A	
6	t_B	t_1 t_2	t_E	some missing				A	
7	$t_B=t_1$ t_2		t_E	all missing	A				
8	$t_B<t_1$ t_2		t_E	all missing					
9	t_B		t_1 $t_2<t_E$	all missing					
10	$t_B=t_1$	$t_2=t_E$		perfect	A	A		A	S
11	$t_B<t_1$	$t_2=t_E$		perfect		A		A	
12		$t_B=t_1$	$t_2<t_E$	perfect	A	A		A	S
13		$t_B=t_1$ $t_2=t_E$		perfect	A	A		A	S
14	$t_B=t_1$	$t_2<t_E$		some missing	A	A		A	
15	$t_B<t_1$	$t_2<t_E$		some missing		A		A	
16		$t_B=t_1$ $t_2<t_E$		some missing	A	A		A	
17		$t_B=t_1$ t_2	t_E	some missing	A	A		A	
18	t_B	t_1 $t_2<t_E$		some missing				A	
19		$t_B<t_1$ $t_2<t_E$		some missing				A	
20		$t_B<t_1$ t_2	t_E	some missing				A	
21		$t_B<t_1$	$t_2<t_E$	some missing				A	
22	$t_B=t_1$ t_2	t_E		all missing	A				
23	$t_B<t_1$ t_2	t_E		all missing					
24		t_B	t_1 $t_2<t_E$	all missing					

The BH procedure uses the complete cases 1, 10, 12, and 13, but does not use the complete cases 2 and 11. It also uses the partial cases 3, 14, 16, and 17, and cases 7 and 22 for which there is no data in interval B. The BI procedure uses all the complete cases, more of the partial cases and none of the cases with no data. We thus think the BI procedure will tend to produce smaller variances than the BH procedure since it uses more of the available data. However, it is not clear in general which of these two procedures has the smaller proportion of needed data that is missing.

The CM procedure is appealing for the restricted universe since it uses all the complete cases (except in the rare situation when there is at least one original sample person present for every month of interval B, but none of them are present for the entire interval), and none of the other cases. It should thus have fairly small variances and has only the slight bias indicated in Section 3. However, it is not applicable to the unrestricted universe.

The HW procedure uses the same complete cases as the BH procedure, except it does not use these cases when the householder is not an original sample person, and it uses none of the other cases. However, it is not applicable to the RM, and most other longitudinal household definitions.

The AW procedure is the most aggressive in utilizing partial data. It uses all the complete and partial cases while avoiding the cases with no data. Also note that it assigns smaller weights, in general, to the partial cases than the complete cases. We believe it will tend to produce the smallest variances for most definitions, particularly in the unrestricted universe, but also tends to have the highest proportion of data that is needed for estimation but unavailable.

REFERENCES

Dicker, Marvin and Casady, Robert J. (1982), "A Reciprocal Rule Model for a Longitudinal Family Unit," American Statistical Association - Proceedings of the Social Statistics Association.

McMillen, David B. and Herriot, Roger A. (1984), "Toward a Longitudinal Definition of Households," SIPP Working Paper Series, No. 402, Bureau of the Census, Washington, D.C.

Nelson, D.D., McMillen, D.B., and Kasprzyk, D. (1984),"An Overview of the Survey of Income and Program Participation", SIPP Working Paper Series, No. 401, Bureau of the Census, Washington, D.C.

Whitmore, R.W., Cox, B.G., and Folsom R.E. (1982), "Family Unit Weighting Methodology for the National Household Survey Component of the National Medical Care Utilization and Expenditure Survey," Research Triangle Institute report.

LONGITUDINAL ITEM IMPUTATION IN A COMPLEX SURVEY

Michael E. Samuhel and Vicki Huggins, Bureau of the Census

INTRODUCTION

Missing data in sample surveys are of two general forms. Unit nonresponse occurs when no information is available to the survey for an entire sample unit, such as a person, or household, or hospital. Some information may, however, be available from other kinds of records such as those used to define the sample frame. The reasons for unit nonresponse vary; for example, a person may refuse to respond, be away from home, or be impossible to locate. Typically, this form of nonresponse is handled in part by a call-back strategy. That is, the interviewer makes repeated attempts to contact the unit. If the call-back strategy fails, or is not feasible, weights can be assigned to the responding units (Cochran, 1977).

The other type of nonresponse is item nonresponse. It occurs when the unit supplies information for some but not all of the variables. For example, a person may answer questions about age, race, and sex but not about income; or the information may be deleted by an edit failure. Depending upon the intended uses of the data, item nonresponse can be handled with two different but overlapping approaches. Either the data can be completed using imputation methods, or the recorded data can be used with modified estimation methods. The modified estimation methods may also be used to impute the missing data.

The focus of this paper is the imputation of categorical data in a longitudinal survey. Statistical research pertaining to missing categorical data has considered censored, discrete random variables and partially or completely unobserved data in contingency tables. Harley's (1958) solution to the problem of estimating the rate parameter for a censored Poisson random variable is a special case of what was later called the EM algorithm. Fuchs (1982) applied the EM algorithm to find maximum likelihood estimates for parameters in a log-linear model, when the values of one or more variables are missing for subsets of the cross-classified data. Chen and Fienberg (1974) developed models for analyzing contingency tables with supplemental marginal totals.

Unfortunately, none of these methods offer solutions to the problem of missing categorical data in complex, longitudinal surveys such as the Survey of Income and Program Participation (SIPP). Although a contingency table could be constructed from monthly responses to a categorical survey item over a year, the resulting twelve dimensional table would be exceedingly sparse. In addition, the application of log-linear models or the EM algorithm to such tables would be computationally difficult.

In this paper we describe a general method for imputing missing categorical items in longitudinal surveys. We show that the longitudinal data, completed according to the method, provides unbiased estimates of the probability of occurrence of the various response patterns, assuming that the data are observed at random and missing at random (Rubin, 1976). The importance of longitudinal information for imputing missing data is discussed, and a statistic measuring the amount of information available is described.

The imputation methodology described here was developed from data collected by the Income Survey Development Program (ISDP). The method is suggested as the fundamental tool for imputing missing, longitudinal, categorical items in the Survey of Income and Program Participation (SIPP). However, its implementation can occur only after further development and modifications. Here, it is described as a general, statistical approach applicable to any longitudinal survey. The data from the ISDP is utilized only to explain the method and provide examples.

The Income Survey Development Program (ISDP) was initiated to gain experience with the data collection and data analysis requirements of SIPP. The ISDP is a longitudinal survey consisting of two national panels (1978, 1979). The sample design is a multi-stage stratified sample of the United States population. Sampling elements are housing units not households (which may move) or persons. The first sampling stage involves the definition of the United States in terms of counties or groups of counties called primary sampling units (PSU's), which are stratified. At the second stage, a sample of addresses within the PSU's is selected. To minimize the inconvenience to sample participants, interviews are conducted every three months. Each household is assigned to one of three rotation groups (A,B,C). Every three months all the households in a rotation group are interviewed and data is collected for each of the previous three months. A wave is the time period during which each rotation group is interviewed once. Data from each wave is published by the United States Bureau of the Census as a cross-sectional file. The longitudinal data for our imputation research is an annual file, constructed by merging five waves of ISDP data from the 1979 panel.

THE IMPUTATION OF MISSING LONGITUDINAL CATEGORICAL SURVEY ITEMS

Many of our activities today are the direct result of events which occurred yesterday. Last night we may have arrived home late, returning from a long trip. Today, it is likely that we will need to stop off at the gas station to refill our car's fuel tank. Or perhaps yesterday we were layed-off from our job. Today we are reading the employment opportunities section of the newspaper.

Analogously, in the ISDP, there are strong dependencies between the monthly values of the survey items. For example, fitting a logistic regression of the receipt of wages and salaries in July on the receipt reported in other months, we found the parameters for the months June, August, and November to be significantly different from zero. Similiar results where obtained in regressions of each month on the remaining months.

Define a longitudinal record for a survey unit to be the set of responses recorded over a fixed time period. In the ISDP as well as SIPP, the survey unit is a household, but other examples of survey units include the person, family, and employer. In this paper, the survey person is the unit of analysis. The set of responses on the longitudinal record may be any combination of survey items. Here, we restrict ourselves to a single item recorded monthly for one year. For example, the receipt of wages and salaries.

The following example illustrates the imputation process. Consider the ISDP survey item indicating whether a person had a job or business during a month. Further, consider the set of individuals in rotation group A who responded "yes" from January thru November 1979, but did not respond in December, 1979. The longitudinal record for these individuals is given by

$$X = (0,0,0,0,0,0,0,0,0,0,0,2),$$

where $X_t = 0$ $(t=1,\ldots,12)$, if the response in the t^{th} month is "yes", $X_t = 1$ if the response is "no", and $X_t = 2$ indicates missing data. Either "0" or "1" is an admissible imputation value for X_{12}. Based on the individuals in rotation group A who reported data in every month from January to December we estimate

$$\text{Prob}\,(X_{12} = 0 \mid X_1 = 0, X_2 = 0,\ldots,X_{11} = 0)$$
$$= \frac{2313}{2379} = 0.9723, \text{ and}$$

$$\text{Prob}\,(X_{12} = 1 \mid X_1 = 0, X_2 = 0,\ldots,X_{11} = 0)$$
$$= 1 - .9723 = 0.0277\,.$$

Generating a random number between zero and one, we impute $X_{12} = 0$ if the random number is less than or equal to 0.9723, otherwise we impute $X_{12} = 1$.

This imputation procedure can be applied to any categorical survey item with any combination of missing months. Consider the sample item indicating the monthly receipt of wages and salaries and the following longitudinal record for persons in rotation group A

$$X = (0,0,0,0,0,0,0,2,2,2,0,0)\,.$$

Based on those persons responding in all twelve months, we estimate

$$\text{Prob}\,(X_8 = x_8, X_9 = x_9, X_{10} = x_{10} \mid \quad (1)$$
$$X_1 = 0,\ldots,X_7 = 0, X_{11} = 0, X_{12} = 0)$$

$$= \frac{1120}{1140} = 0.9823 \text{ if } X_8 = 0,\ X_9 = 0,\ X_{10} = 0,$$

$$= \frac{10}{1140} = 0.0088 \text{ if } X_8 = 1,\ X_9 = 0,\ X_{10} = 0,$$

$$= \frac{4}{1140} = 0.0035 \text{ if } X_8 = 0,\ X_9 = 0,\ X_{10} = 1,$$

$$= \frac{3}{1140} = 0.0026 \text{ if } X_8 = 1,\ X_9 = 1,\ X_{10} = 0,$$

$$= \frac{2}{1140} = 0.0018 \text{ if } X_8 = 0,\ X_9 = 1,\ X_{10} = 0,$$

$$= \frac{1}{1140} = 0.0009 \text{ if } X_8 = 1,\ X_9 = 0,\ X_{10} = 1.$$

Here, we impute the entire subvector (x_8, x_9, x_{10}) based on a random draw from a uniform (0,1) distribution.

The imputation process is formalized by letting the random variable X represent the responses (and missing data) on a longitudinal record. The vector X = x can be partitioned into subvectors x_m and x_r, representing the missing and recorded monthly valves, respectively. On the i^{th} longitudinal record, we impute the missing items X_{mi} based on the reported values x_{ri}. The imputed values are a random draw from the conditional distribution $f(x_m \mid X_r = x_{ri})$, emperically estimated from the longitudinal records with values reported in every month.

AN UNBIASED ESTIMATE OF THE OCCURRENCE PROBABILITY OF A LONGITUDINAL PATTERN

Response patterns to survey items are singularly important in longitudinal surveys. The longitudinal data is collected so that changes over time of the survey items can be analyzed. For example, a researcher may wish to accurately estimate the average duration of unemployment or the length of time an individual participates in a social welfare program. It is important that the imputations do not disrupt the frequency distribution of response patterns and bias these longitudinal estimates.

Consider a simple random sample of a size n without nonresponse. The longitudinal records for individuals in the labor force every month are represented by

$$X = (0,0,0,0,0,0,0,0,0,0,0,0) \quad (2)$$

Let the binomial random variable T represent the number of times the pattern (2) occurs. It follows that

$$\frac{1}{n}\,T\,(X_1 = 0, X_2 = 0,\ldots,X_{12} = 0) \quad (3)$$

is an unbiased estimate of

$$\text{Prob}\,(X_1 = 1, X_2 = 1,\ldots,X_{12} = 1).$$

Of course, in longitudinal surveys with complex sample designs like SIPP, the statistic (3) would need to be modified to reflect the particular survey design.

In longitudinal files, completed according to the imputation method described above, statistics analagous to (3) are also unbiased estimates of the probability that the particular pattern occurs; provided the data are missing at random and observed at random (Rubin, 1976). We prove this result for longitudinal records containing two time periods. Without loss of generality the result extends to longitudinal records of any length.

THEOREM

Consider the longitudinal record $(X_1 = a, X_2 = b)$, where a and b represent the only values of the categorical random variables X_1 and X_2. In a simple random sample of size n, completed by imputation, let the binomial random variable $T'(X_1 = a, X_2 = b)$ represent the number of occurrences of the longitudinal record. Assuming the data are observed at random and missing at random.

$$\frac{1}{n}\,T'(x_1 = a, x_2 = b)$$

is an unbiased estimate of

$$\text{Prob}(X_1 = a, X_2 = b)$$

Proof:

The pattern (X_1 = a, X_2 = b) can arise in the imputed sample in four ways:

1) (X_1 = a, X_2 = b) is reported,
2) X_1 = a is imputed given X_2 = b is reported
3) X_2 = b is imputed given X_1 = a is reported,
4) (X_1 = a, X_2 = b) is imputed.

Define the binomial random variable T( ) as the number of occurrences of the event in parentheses. For example, using an astrisk to indicate imputed counts,

$$T^*(X_1 = a \mid X_2 = b)$$

represents the number of times that X_1 = a is imputed given that X_2 = b is reported.

The total number of times the pattern

$$(X_1 = a, X_2 = b)$$

occurs in the sample, completed by imputation, can be decomposed into terms corresponding to the four ways the pattern (a, b) arises,

$$\begin{aligned} T'(X_1 = a, X_2 = b) = \; & T(X_1 = a, X_2 = b) + \\ & T^*(X_1 = a \mid X_2 = b) + T^*(X_2 = b \mid X_1 = a) + \\ & T^*(X_1 = a, X_2 = b). \end{aligned} \tag{4}$$

Let the indicator vector Y = (Y_1, Y_2) represent the reporting status of the elements in the longitudinal record. That is,

$$\begin{aligned} Y_i &= 1 \text{ if } X_i \text{ is reported (i=1,2),} \\ &= 0 \text{ otherwise .} \end{aligned}$$

The expected value of the sum (4) with respect to the data reported in the sample is

$$\begin{aligned} E(T'(X_1 = a, X_2 = b) \mid T(X_1 = a, X_2 = b)) = \; & \\ T(X_1 = a, X_2 = b, Y_1 = 1, Y_2 = 1) + & \\ T(X_2 = b, Y_1 = 0, Y_2 = 1) \cdot & \\ \left[\frac{T(X_1 = a, X_2 = b, Y_1 = 1, Y_2 = 1)}{T(X_2 = b, Y_1 = 1, Y_2 = 1)}\right] + & \\ T(X_1 = a, Y_1 = 1, Y_2 = 0) \cdot & \\ \left[\frac{T(X_1 = a, X_2 = b, Y_1 = 1, Y_2 = 1)}{T(X_1 = a, Y_1 = 1, Y_2 = 1)}\right] + & \\ + T(Y_1 = 0, Y_2 = 0) \cdot & \\ \left[\frac{T(X_1 = a, X_2 = b, Y_1 = 1, Y_2 = 1)}{T(Y_1 = 1, Y_2 = 1)}\right]. & \end{aligned} \tag{5}$$

Note that the random variables in the conditional expectation (5) are multimonial. The expectation with respect to all possible samples is found by applying the following result.

LEMMA

Let (X_1,...,X_k) be multimonial (n;P_1,...,P_k) random variables, then X_1 and X_3 are independent given $X_1 + X_2 = \ell$ and

$$E\; X_3 \frac{X_1}{X_1 + X_2} = nP_3 \frac{P_1}{P_1 + P_2} .$$

The expectation of (5) with respect to all possible samples follows from the lemma. In addition, the assumption that the data are observed at random and missing at random asserts the independence of the indicator random vector Y and the random variables in the longitudinal record.

$$\begin{aligned} & E(T'(X_1 = a, X_2 = b)) = \\ & E_2\, E_1\, (T'(X_1 = a, X_2 = b) \mid T(X_1 = a, X_2 = b)) = \\ & n\,\text{Prob}(X_1 = a, X_2 = b)\,\text{Prob}(Y_1 = 1, Y_2 = 1) + \\ & n\,\text{Prob}(X_2 = b)\,\text{Prob}(Y_1 = 0, Y_2 = 1) \cdot \\ & \left[\frac{\text{Prob}(X_1 = a, X_2 = b)}{\text{Prob}(X_1 = a, X_2 = b) + \text{Prob}(X_1 = b, X_2 = b)}\right] + \\ & n\,\text{Prob}(X_1 = a)\,\text{Prob}(Y_1 = 1, Y_2 = 0) \cdot \\ & \left[\frac{\text{Prob}(X_1 = a, X_2 = b)}{\text{Prob}(X_1 = a, X_2 = a) + \text{Prob}(X_1 = a, X_2 = b)}\right] + \\ & n\,\text{Prob}(Y_1 = 0, Y_2 = 0)\;\; \text{Prob}(X_1 = a, X_2 = b) \\ & = n\,\text{Prob}(X_1 = a, X_2 = b) \end{aligned}$$

QED

The theorem is extended to longitudinal records of any length by adding the appropriate terms to equation (3).

THE EXPECTED NUMBER OF INCORRECT IMPUTATIONS

Longitudinal data by itself may not always be sufficient to accurately impute missing data. The amount of information available longitudinally can be measured by estimating the expected number of incorrect imputations. Consider the longitudinal record for the monthly receipt of wages and salaries,

$$X = (0,0,0,0,0,0,0,0,0,0,0,2), \tag{6}$$

where X_t = 0 indicates receipt and X_t = 2 (t=1,...,12) indicates missing data. The probability

$$\text{Prob}(X_{12} = 1 \mid X_1 = 0,\ldots,X_{11} = 0)$$

is estimated from the completely reported cases as

8/1236 = 0.0065. This probability is independent of but equal to the probability of imputing $X_{12} = 1$. Consequently, the probability that $X_{12} = 1$ is imputed and is correct is $(0.0065)^2$. Similiarly, the probability that $X_{12} = 0$ is imputed and is correct is $(0.9935)^2$. It follows that the estimated probability of an incorrect imputation for the longitudinal record (6) is

$$1 - (0.0065)^2 + (0.9935)^2 = 0.0013.$$

Since, there are seventeen individuals in the file with this longitudinal record, it follows that the estimated number of incorrect imputations is 17(0.013) = 0.22.

The need to include demographic information would be indicated by an estimated number of incorrect imputations greater than some predetermined value. Consider again the longitudinal record for the monthly receipt of wages and salaries,

$$X = (0,0,0,0,0,0,0,2,2,2,0,0).$$

Thirty-eight persons in rotation group A had this pattern. Using the probabilities given in (1), the estimated expected number of incorrect imputation is

$$38(1-0.9825^2 - 0.0088^2 - 0.0035^2 - 0.0026^2 - 0.0018^2 - 0.0009^2) = 1.25 .$$

Here, we want to use demographic information to choose the most appropriate donor pattern. One approach is to include associated survey items as elements in the longitudinal record. For example, to impute the monthly receipt of wages and salaries, we can include in the longitudinal records survey items indicating seasonal or part time workers. A logistic model may also be useful, especially when the data are sparse. Letting the polychotomous variable Y represent the available donor patterns, we can regress Y on concomitant data, represented by the vector X. Based on the concomitant information, the probability of pattern h for the i^{th} longitudinal record is

$$\text{Prob}(Y_i = h) = \frac{e^{B_h X_i}}{1 + e^{B_h X_i}} .$$

The pattern selected for imputation can be the one with the highest probability, or the decision can be based on a random number generated between zero and one.

CODING PATTERNS

The responses on any longitudinal record can be summarized as a single number. Consider the longitudinal record

$$X = (0,0,0,0,0,0,1,1,1,2,2,2),$$

representing the receipt of wages and salaries from January (X_1) to December (X_{12}). This pattern can be represented in base ten as

$$377 = (2x3^0) + (2x3^1) + (2x3^2) + 3^3 + 3^4 + 3^5 .$$

In general, any pattern in an annual file of monthly categorical data can be represented by the polynomial

$$P_\ell = \sum_{k=1}^{12} c_k B^{k-1} ,$$

Each pattern has a unique base ten representation, because the transformation is one-to-one and onto, the index k represents the months in the longitudinal file in reverse order. That is, k=1 represents December, k=2 represents November, and so on. The coefficients c_k represent the monthly values of the item. The letter B represents the appropriate base. Typically, the base is one more that the highest coefficent (c_k).

Coding the longitudinal record patterns as base ten numbers operationally simplifies the imputation process. Consider the longitudinal record

$$X = (0,0,0,0,0,0,0,0,0,2,2,2),$$

indicating the receipt of wages and salaries in each month from January thru December. The receipt of wages and salaries in the t^{th} month is denoted by $X_t = 0$, and a missing monthly item is denoted by $X_t = 2$. This pattern is represented in base ten by the number 26. Because the transformation to base ten is unique, all individuals in the data file with the value 26 for their pattern have reported the receipt of wages and salaries from January to September, but did not respond to the item from October to December.

Donor patterns from the cases, reporting values in every month, are identified by subtraction. For example, the donor pattern.

$$(X_{10} = 0, X_{11} = 0, X_{12} = 0)$$

is identified by subtracting the base three number 222 from the longitudinal pattern

$$\begin{array}{r} 000000000222 \\ -222 \\ \hline 000000000000 \end{array}$$

The equivalent operation could also be done in base ten. Noting that 222 is represented in base ten by 26, the donor pattern ($X_9 = 0, X_{10} = 0, X_{11} = 0$) is the base three representation of (26-26) = 0. Similiarly, all possible donor patterns i.e., 000 thru 111 and found by subtracting from 26 the corresponding base ten numbers 26 through 0.

APPLICATIONS AND EXTENSIONS OF THE METHOD

Limitations on the number of pages available in these proceedings preclude a complete discussion of our research on longitudinal item imputation. A more extensive description, especially as it applies to the Survey of Income and Program Participation, can be found in Samuhel and Huggins (1984).

Acknowledgements

The author wishes to thank Dr. Patrick Kelly for many helpful discussions and acknowledge his contributions to the proof of the theorem.

References

Chen, T. and Fienberg, S.E. 1974. Two dimensional contingency tables with both completely and partially classified data. Biometrics 30: 629-642.

Cochran, W.G. 1977. Sampling Techniques, New York: John Wiley and Sons, Inc.

Fuchs, C. 1982. Maximum likelihood estimation and model selection in contingency tables with missing data. Journal of the American Statistical Association 77: 270-278.

Hartley, H.O. 1958. Maximum likelihood estimates from incomplete data. Biometrics 14: 174-194.

Rubin, D.B. 1976. Inference and Missing data. Biometrika 63: 581-592.

Samuhel, M.E. and Huggins, V.J. 1984. Longitudinal item imputation in a complex survey. Survey of Income and Program Participation Working Paper Series. United States Department of Commerce.

EARLY INDICATIONS OF ITEM NONRESPONSE ON THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

by John F. Coder and Angela M. Feldman

Introduction

The Survey of Income and Program Participation (SIPP) promises to become the most important source of data for measuring the level of and changes in the economic well-being of the U.S. population. Collection of these data began in the fall of 1983. The survey design for the initial sample of 25,900 housing units in the noninstitutional population, calls for each household to be interviewed at 4-month intervals over a 2-1/2 year period. The sample is divided into 4 rotations or panels of equal size and one panel is interviewed in each month throughout this period resulting in a total of eight personal contacts by Census interviewers for each sample household.

The first interviews in this new survey were conducted during October, November, and December of 1983, and January 1984. The questionnaire used to collect information in the initial interview concentrates on labor force participation and sources and amounts of income. Most data is recorded separately by month for the 4-month reference period ending in the month prior to the month of interview. For example, data collected in the October 1983 interviews covered the June through September period. Most interviews were completed during the first 2 weeks of the interview month.

The primary purpose of this paper is to present some preliminary indications of the item nonresponse rates for the first interviews of SIPP. These rates of nonresponse cover labor force, income recipiency, and income amounts. The effect of self or proxy respondents on nonresponse rates is discussed for a selected group of items. Some data on other aspects of the survey have also been included. These are overall household noninterview rates, average times required for interviews, and use of callback procedures to obtain missing information.

Item Nonresponse

Item nonresponse is defined in this paper to mean a missing answer to a specific question that should have been answered. Item nonresponse can result for many reasons, the most frequent being lack of knowledge by the respondent, i.e., "Don't Knows," and refusals to answer. Nonresponse can also result when the interviewer fails to record a response in the correct location or follows an incorrect path within the questionnaire design.

Labor Force Items--Table 1 shows preliminary nonresponse rates for items 2a, 2b, 4, 5a, 5b, 6a, 6b, 6c, 7a, 7b, and 8a of the labor force and recipiency section on the first interview questionnaire. The questions themselves are shown in Figure 1.

In general, the nonresponse rates for the labor force questions were low (see table 1). The nonresponse rate on item 2a, incidence of looking for work or on layoff for persons who did not work at all during the reference period (nonworkers) was only 0.4 percent. About 6.7 percent of the nonworkers reporting looking or on layoff had a nonresponse for item 2b, the number of weeks spent looking or on layoff. The comparable nonresponse rates for workers were 1.0 percent for incidence of looking or on layoff (item 7a) and 3.2 percent for item 7b, the number of weeks spent looking or on layoff. The nonresponse rate for item 4, asking if the respondent held a job or business during the entire 4-month reference period, was less than 0.1 percent.

One of the questions with a relatively high nonresponse rate in the labor force section was item 5b covering the number of weeks absent without pay for persons having a job for the entire period. The nonresponse rate for this question was 11.6 percent.

Item 8a is the question covering the number of hours usually worked per week during the 4-month period. This critical data item was missing for 1.3 percent of the 25,510 sample persons reporting a job or business during the reference period.

Income Recipiency.--A major portion of the questionnaire was designed to determine the sources of income received during the 4-month period by each household member age 15 years old and over. A total of 52 different income sources (other than earnings from employment) were covered in the survey. Tables 2 and 3 show income recipiency nonresponse rates and ratios of nonresponses to "YES" responses for SIPP and the March 1983 CPS for a selected group of income types. The rates refer to the 4-month reference period for SIPP and calendar year 1982 for the March CPS.

The nonresponse rates for SIPP are extremely low and vary only slightly by rotation. The nonresponse rate on recipiency for SIPP ranged from less than 0.1 for Aid to Families with Dependent Children and private pensions to 1.3 percent for stocks or mutual funds. In contrast, the rates for the March 1983 CPS clustered around the 10-percent level. These rates for the March CPS are largely attributable to the 7 percent household noninterview rate on the income supplement questionnaire.

The last two columns of table 3 show the ratios of nonresponses to "YES" responses for SIPP and the March CPS. This measure of nonresponse may be better than the overall nonresponse rate because it provides a measure that is relative to the size of the recipient universe. The March CPS ratios are again much higher than those encountered in the first interview of SIPP. This difference is also related to the 7 percent March supplement noninterview rate. Given this fixed nonresponse rate the ratio is inversely related to the proportion of the population receiving a specific income type. This is evident by the large ratio of 4.01 for Aid to Families with Dependent Children. The ratio itself means that, in this case, the number of nonresponses and, therefore, imputations required exceeded the number of "YES's" by a factor of 4 to 1.

Hourly Wage Rates.--The nonresponse rates on hourly wages are shown in table 4. These rates are shown separately by type of respondent. The nonresponse rate was 9.5 percent overall, 5.1 percent for self response and 16.7 percent for proxy response. The overall nonresponse rate for hourly wages increased from the 7.8 percent level in October to 10.5 percent in January. This resulted mainly from an increase in the nonresponse rate for proxy responses of from 13.8 percent in October to 19.2 percent in January. Approximately 62 percent of the respondents were "self."

Monthly Wage or Salary Income.--Table 5 contains the nonresponse rates for the monthly amounts of wage and salary income. The nonresponse rate overall averaged about 6.2 percent for the initial SIPP interviews. The rate for self respondents, which accounted for 64 percent of the total, was lower, 4.6 percent, while the rate for proxy respondents was 9.0 percent. The 9.0-percent nonresponse rate for proxy interviews on monthly earnings amounts was considerably lower than the comparable rate of 16.7 percent for hourly wage amounts. Nonresponse rates increased from 5.4 percent to 6.7 percent between October and January.

Self-Employment Income.--Nonresponse rates for self-employment income have traditionally exceeded those for most income types. The items in the self-employment section of the SIPP questionnaire cover monthly amounts of "salary" and other income received by owners of businesses, professional practices, farms, etc. The question is not designed to obtain estimates of the business's net profit on a monthly accounting period. An additional question was included covering estimated net profit for the entire 4-month reference period. The nonresponse rate overall for the monthly salary or other income received by the self-employed was 14.0 percent (see table 6). The nonresponse rate for proxy interviews exceeded that of self-responses by a considerable margin. The rate for proxy interviews was 22.3 percent compared to 9.8 percent for self responses. The October nonresponse rate of 13.6 percent was not significantly different from the January rate of 15.1 percent. About two-thirds of respondents for this item were "self."

Interest Income.--Table 7 contains nonresponse rates for interest amounts received during the SIPP 4-month reference period. These rates cover the interest amount received from one or more of the following sources: 1) regular or passbook savings, 2) money market deposit accounts, 3) certificates of deposit, or other savings certificates, and 4) NOW accounts or other interest earning checking accounts. The nonresponse rate for interest income from these sources was 34.6 percent. The rate in January was 35.4 percent, somewhat higher than the 32.6 percent for October. About 4 percent of the total number of nonresponses on interest amounts can be attributed to refusals. The remainder were mainly categorized as "Don't Knows." A "Don't Know" response to interest income was followed by a question to obtain the balance or amount in the account. The nonresponse rates for this item are also shown in table 7. The nonresponse rate for balances in savings was 24.2 percent. In combination these two nonresponse rates indicate that both the interest amount and the balance amount were missing in only about 13.3 percent of the sample cases for these sources of interest income.

Dividend Income.--The questions covering the amount of dividend income received were divided into two categories, those dividends actually received and those credited against a margin account or automatically reinvested in additional shares of stock. As indicated by the data in table 8, the nonresponse rates for these two categories differ significantly. The rate for dividends actually received was 9.4 percent. The rate for dividends credited was 30.7 percent.

Noninterview Rates

The noninterview rate is a measure of the proportion of occupied housing units, i.e., those eligible for interview, for which interviews were not obtained. As mentioned earlier the total sample size for the 1983 SIPP was about 25,900 housing units. Of this total about 4,600 were not eligible for interview. These ineligible units were found to be vacant, demolished, under construction, or unoccupied for other reasons. This left 19,900 households eligible to be contacted. Interviews were not obtained for 4.8 percent of this group (see table 9). Most noninterviews, about 77 percent, were refusals to participate. The remainder of the total noninterview rate consisted of situations classified as "no one home" and "temporarily absent." These classifications were assigned after repeated visits failed to yield a contact.

The noninterview rate varied considerably by region of the Country. The lowest noninterview rate was 2.4 percent from the Kansas City Regional Office that covers Kansas, Missouri, Iowa, Minnesota, and Wisconsin. The highest noninterview rate was 10.1 percent from the New York Regional Office covering the parts of New York and New Jersey in the vicinity of New York City.

There was slight variation in the noninterview rates by month of interview, however, these rates were not significantly different from one another. The rate for the first month of interview was 5.1 percent compared to 4.3 percent, 5.2 percent, and 4.8 percent in the succeeding 3 months, respectively. The overall noninterview rate for SIPP (4.8 percent) was not significantly different from the overall rate for the March 1983 CPS (4.4 percent) or the rate for the panel coming into the March 1983 CPS for the first time (5.4 percent). As noted earlier, about 7.0 percent of the March CPS sample households completed the monthly labor force questions but were noninterviews on the income supplement. These cases are in addition to the 4.4 percent household noninterviews.

Callback Items

The design of the SIPP questionnaire incorporated procedures for following up on missing responses to items identified as either especially important to the overall quality of the survey data or with previously noted high nonresponse rates. The first step in this process was the determination that the answer to the designated question would be available from another household member not present at the time of the interview or at a later date. If so, the interviewers, in most cases, called back by telephone to obtain the missing information. The data in table 10 summarize use of the callback system.

The callback system appears to be most effective for obtaining missing data on amounts of monthly wage and salary income. About 600 cases were marked for callback for these amounts. The procedure obtained responses to the missing earnings amounts in about 7-out-of-10 cases. Use of the callback was less successful in obtaining missing amounts for the other income sources. Slightly more than half (54 percent) of the callbacks were successful for obtaining data for the monthly amount of salary and other income received from self-employment. Attempts to follow up on amounts of interest and dividend income from various sources proved to be even less effective. About 45 percent of the respondents were able to supply an amount when contacted by an interviewer. Use of the callback procedures appears to have declined between the October and January interviews. Generally, the number of cases marked for follow-up in January were lower than October. While less frequent use of the callback might have been related to a reduced need for follow-up, nonresponse rates for some of these income types tended to increase between October and January, indicating the opposite.

Interview Time

The time required to conduct an initial SIPP interview is potentially quite long given the number of questions. Obviously households with a large number of adult members, those 15 years old and over, are those that are exposed to the longest overall interview times, on average. The data in table 11 provide the first estimates of interview times based directly on times entered on each person's questionniare by the interviewers. The time required to complete the household control card and roster was added to the interview time on the first questionnaire for the household. These estimates are shown by size of household for the first interview period of SIPP.

The median interview time was 43 minutes for all households in the first interview. The median interview time declined steadily from 48 minutes in October to 41 minutes in January. The median household interview time for 1-person households was about one-half hour while that for 4-person households was one hour and ten minutes. Households with 5, 6, and 7 or more members required proportionally more time for interviews.

Summary

This examination of some of the early "returns" from the 1983 SIPP are, for the most part, encouraging. The household noninterview rate was lower than most had anticipated. The item nonresponse rates were much lower than those experienced in the March CPS. Proxy responses caused significantly higher nonresponse rates for some of the key items studied.

There is reason for concern, however, in several areas and these should be watched closely. The first is the general trend toward higher nonresponse rates between October and January interviews. The second is the relatively high noninterview rate for the New York area. While this is consistent with our experiences in other surveys, this rate should be monitored closely as will the rates in the other regions.

The next step in the evaluation of the 1983 SIPP data will be comparison of the survey estimates of income recipients with figures derived from program statistics and other independent sources. This analysis will provide a very important look at the magnitude of survey underreporting, a major concern of SIPP and other household income surveys.

Figure 1. Selected Labor Force Questions

NONWORKERS

2a. Even though ... did not have a job during this period, did ... spend any time looking for work or on layoff from a job?

|__| YES -- ASK 2b

|__| NO

2b. In which weeks was ... looking for work or on layoff from a job?

WORKERS

4. Did ... have a job or business, either full or part time, during EACH of the weeks in this period?

|__| YES -- ASK 5a

|__| NO -- ASK 6a

5a. Was ... absent without pay from ...'s job or business for any FULL weeks during the 4-month period?

|__| YES -- ASK 5b

|__| NO

5b. In which weeks was ... absent without pay?

WORKERS WITH WEEKS WITHOUT A JOB OR BUSINESS

6a. In which weeks did ... have a job or business?

6b. Was ... absent from work for any full weeks without pay?

|__| YES --ASK 6c

|__| NO

6c. In which weeks was ... absent without pay?

7a. During the weeks that ... did not have a job did ... spend any time looking for work or on layoff?

|__| YES -- ASK 7b

|__| NO

7b. In which of these weeks was ... looking for work or on layoff from a job?

WORKERS

8a. In the weeks that ... worked during the 4-month period, how many hours did ... usually work per week?

Table 1. Selected Item Nonresponse Rates for the Labor Force Items on the 1983 SIPP: Interview No. 1

Item	Total	Rotation			
		One	Two	Three	Four
2a	0.4	0.4	0.4	0.4	0.3
2b	6.7	8.2	6.8	5.9	5.9
4	0.1	0.1	0.1	(Z)	0.1
5a	0.1	0.1	0.1	0.1	0.1
5b	11.6	12.6	11.0	8.2	14.4
6a	2.2	2.9	2.0	1.9	1.8
6b	3.3	6.6	2.3	1.8	1.4
6c	6.8	2.1	12.2	3.3	10.5
7a	1.0	0.9	1.0	1.1	0.9
7b	3.2	4.7	3.7	2.0	2.0
8a	1.3	1.3	1.3	1.2	1.2

Z Less than .05 percent.

Table 2. Selected Item Nonresponse Rates for Income Recipiency During the 4-month Reference Period on the 1983 SIPP: Interview No. 1

Income type	Total	Rotation			
		One	Two	Three	Four
Social Security.....	0.6	0.6	0.6	0.5	0.5
Unemployment compensation.......	0.1	0.1	0.1	0.1	0.1
Veteran's payments..	0.2	0.2	0.2	0.2	0.2
Aid to Families with Dependent Children.	(Z)	(Z)	(Z)	(Z)	(Z)
Food stamps.........	0.3	0.4	0.4	0.2	0.2
Private pensions....	(Z)	(Z)	(Z)	0.1	(Z)
Savings accounts....	1.0	0.8	0.8	1.1	1.1
Shares of stock or mutual funds.......	1.3	1.4	1.1	1.3	1.3
Rental property.....	1.0	1.0	0.7	0.9	1.1

Z Less than .05 percent.

Table 3. Selected Income Nonresponse Rates from the March 1983 CPS, Ratio of Nonresponses to "YES" Responses for the March 1983 CPS, and Ratio of Nonresponses to "YES" Responses for Interview No. 1 of the 1983 SIPP

Income type	March 1983 CPS nonresponse rate	March 1983 CPS ratio of nonresponses to "YES's"	1983 SIPP ratio of nonresponses to "YES's"
Social Security...	9.6	0.61	.03
Unemployment compensation.....	9.6	1.16	.03
Veteran's payments	9.6	1.14	.10
Aid to Families with Dependent Children.........	9.7	4.28	.01
Food stamps.......	6.4	0.84	.07
Private pensions..	9.6	1.64	.01
Savings accounts..	10.4	.21	.02
Shares of stock or mutual funds.....	9.7	0.69	.09
Rental property...	9.7	0.66	.13

Table 4. Nonresponse Rates on Hourly Wage Rate by Type of Respondent for the 1983 SIPP: Interview No. 1

Type of respondent	Total	Rotation			
		One	Two	Three	Four
Total........	9.5	7.8	9.3	10.4	10.5
Self.........	5.1	4.1	4.7	5.9	5.6
Proxy........	16.7	13.8	16.1	18.0	19.2
Proportion of Self Responses...	.62	.62	.60	.63	.64

Table 5. Nonresponse Rates on Monthly Wage and Salary Income by Type of Respondent for the 1983 SIPP: Interview No. 1

Type of respondent	Total	Rotation			
		One	Two	Three	Four
Total........	6.2	5.4	5.8	6.8	6.7
Self.........	4.6	4.2	4.3	4.9	4.9
Proxy........	9.0	7.6	8.4	10.2	10.1
Proportion of Self Responses..	.64	.63	.63	.64	.65

Table 6. Nonresponse Rates on Monthly Amounts of Self-Employment Income for the 1983 SIPP: Interview No. 1

Type of respondent	Total	Rotation			
		One	Two	Three	Four
Total.........	14.0	13.6	12.6	14.6	15.1
Self..........	9.8	9.5	9.7	9.6	10.2
Proxy.........	22.3	21.4	18.6	24.3	24.7
Proportion of Self Responses....	.66	.65	.67	.66	.66

Table 7. Nonresponse Rates for Amounts of Interest Income from the 1983 SIPP: Interview No. 1

Item	Total	Rotation			
		One	Two	Three	Four
Interest amount....	34.6	32.6	33.8	37.1	35.4
Percent refusals.	4.2	4.1	4.0	4.6	4.1
Balance amount.....	24.2	23.6	24.1	24.9	24.1

Table 8. Nonresponse Rates for Amounts of Dividend Income for the 1983 SIPP: Interview No. 1

Item	Total	Rotation			
		One	Two	Three	Four
Dividends received.	9.4	10.3	8.3	9.8	9.3
Dividends credited.	30.7	28.2	33.8	30.1	30.5

Table 9. Household Noninterview Rates by Regional Office for the 1983 SIPP: Interview No.1

Item	Total	Rotation			
		One	Two	Three	Four
Total.........	4.8	5.1	4.3	5.2	4.8
Boston........	3.8	2.9	2.5	5.4	4.6
New York......	10.1	13.3	8.3	10.8	8.4
Philadelphia..	3.0	2.0	3.4	2.5	4.1
Detroit.......	4.1	3.0	3.6	5.4	4.1
Chicago.......	4.8	5.0	3.4	5.7	5.0
Kansas City...	2.4	1.6	1.6	4.0	2.5
Seattle.......	4.7	5.1	4.4	5.2	4.3
Charlotte.....	3.5	4.3	2.7	2.8	3.8
Atlanta.......	4.9	5.4	5.0	5.2	4.2
Dallas........	5.1	5.0	5.1	4.6	5.8
Denver........	5.3	6.1	5.7	4.1	5.5
Los Angeles...	7.5	9.3	6.2	8.9	5.8

Table 10. Success Rates of Callback Items

Item	Total	Rotation			
		One	Two	Three	Four
Success Rates					
Wages and salary....	71.0	76.2	76.9	70.0	59.0
Self-employment.....	54.0	58.6	55.0	48.3	54.5
Interest and dividends..........	44.8	48.4	49.6	38.2	40.8
Number of Callbacks					
Wages and salary....	599	172	143	150	134
Self-employment.....	100	29	20	29	22
Interest and dividends..........	582	192	139	131	120

Table 11. Median Household Interview Times by Number of Members 15 Years Old and Over from the 1983 SIPP: Interview No. 1

Number of persons	Total	Rotation			
		One	Two	Three	Four
Total......	43	48	44	42	41
One...........	29	33	30	26	26
Two...........	44	50	45	42	41
Three.........	57	64	57	55	55
Four..........	70	76	72	67	66
Five..........	83	90	81	84	77
Six...........	98	105	111	101	71
Seven or more..	113	114	(B)	120	94

B Less than 10 sample households.

DISCUSSION

Roy Whitmore, Research Triangle Institute

1. Introduction

The Bureau of the Census is to be commended for presenting papers dealing with proposed methodology during the planning stages of the Survey of Income and Program Participation (SIPP). Presentation of these papers is sure to stimulate constructive suggestions from the scientific community. On the other hand, the SIPP has been in progress since October 1983 and a second panel is to be fielded in January 1985. Hence, it is important for methodological issues that impact directly upon data collection techniques be resolved as quickly as possible.

The first of the three papers being reviewed discusses person-level and household-level cross-sectional weighting procedures for the 1979 Research Panel of the Income Survey Development Program (ISDP), a nationwide field test for the SIPP. The next two papers discuss person-level and household- or family-level longitudinal weighting methods being considered for the SIPP. Each paper will be discussed individually, although it will be noted that some comments pertain to all three papers.

2. Cross-Sectional Estimates for the ISDP by Huang

2.1 Introduction

The following motivational statement is found early in Huang's paper:

"There is a great deal of interest in developing cross-sectional weights at the time of each interview wave."

Due to the use of three rotation groups within each wave, I question the use of wave-specific cross-sectional weights for direct data analysis. The inferential population would be difficult to define due to the sequence of referece periods applicable to the rotation groups (see Figure 1). Wave-specific cross-sectional weights are important for defining longitudinal weights, as is apparent from the other two papers being reviewed. Huang presents his weight formulas in the context of weights for cross-sectional estimates that are time-specific rather than wave-specific, which is probably more useful for data analysis. The formulas presented in the paper can actually be considered to be either wave-specific or time-specific weights.

The proposed weighting procedures are discussed in terms of estimation of the population total

$$X_t = \sum_{i=1}^{N_t} X_t(i), \qquad (1)$$

where $i=1,\dots,N_t$ indexes the "units (persons or households)" in the target population at time t. The weighting formulas presented are for cross-sectional household weights applicable as of either time t or wave w. Since all "adult" members of sample households are interviewed, the cross-sectional household weights can be assigned to all household members for cross-sectional person-level analyses.

Cross-sectional household weights are presented for both the area frame and the list frame samples of the 1979 Research Panel. The proposed weighting procedures are discussed below or each sampling frame.

2.2 Area Frame Sample Weights

The population of inferential interest for the 1979 Research Panel was defined to be the 1979 civilian, noninstitutionalized United States adult population. Standard area frame household sampling procedures were used to select a sample of members of this population in the Wave 1 sample, which was fielded early in 1979. However, only adults (aged $\geq$ 16) in the Wave 1 sample were followed when they moved to new addresses during 1979. Thus, "additional" people who entered the target population during 1979 (by birth, by entering the United States, or by leaving the military or an institution) where only interviewed while living in a household that contained at least one Wave 1 sample member. As a result, the sample fails to adequately reflect "additional" people in the target population. This issue will arise again and be discussed more fully with regard to the two SIPP methodology papers.

Two unbiased cross-sectional time t estimators were discussed for estimation of the population total, X_t. I found it instructive to reformulate these estimators as follows:

$$1. \quad \hat{X}_{t,M} = \sum_{i=1}^{N_t} W_{t,M}(i)\, X_t(i), \qquad (2)$$

where

$$W_{t,M}(i) = \begin{cases} \sum_{j \in S_o} (r_i\, \pi_j)^{-1} & \text{for the i-th household in the time t sample,} \\ 0 & \text{for households not in the time t sample,} \end{cases}$$

S_o = {Wave 1 (time t_o)sample households},

r_i = Number of households in the Wave 1 universe contributing members to the i-th time t household, and

π_j = Selection probability for the j-th Wave 1 sample household.

$$2. \quad \hat{X}_{t,F} = \sum_{i=1}^{N_t} W_{t,F}(i)\, X_t(i), \qquad (3)$$

where

$$W_{t,F}(i) = \begin{cases} \sum_{j \in S_o} S_{ij}(S_{io}\, \pi_j)^{-1} & \text{for the i-th household in the time t sample} \\ 0 & \text{for households not in the time t sample,} \end{cases}$$

S_{ij} = Number of members of the j-th Wave 1 household who belong to the i-th time t household, and

S_{io} = Number of members of the Wave 1 universe who belong to the i-th time t household.

In Huang's paper the estimator (2) is referred to as the multiplicity estimator and (3) is referred to as the fair share estimator. In fact, both (2) and (3) are multiplicity estimators. The difference is that the weight $W_{t,M}$ is based upon household-level multiplicity and $W_{t,F}$ is based upon person-level multiplicity. The paper shows that both estimators provide unbiased estimates of the population total, X_t, invoking the "fair share assumption" for the estimator (3). In fact, the weights $W_{t,M}$ and $W_{t,F}$ are identical to the initial family weights for the national household survey component of the National Medical Care Utilization and Expenditure Survey (NMCUES) [See Whitmore, et al (1982a)]. In the NMCUES report, it is shown that both of these estimators provide unbiased estimates, even without the "fair share assumption" for the estimator (3). Huang's conclusion that the estimator (3) is preferable mainly because it produces less variable weights and hence smaller sampling variances is also supported.

2.3 List Frame Sample Weights

Huang's paper defines the population of inferential interest for the sample based upon SSI and BEOG lists as follows:

"At any time t, the target population consists of the original list frame subpopulation (Groups I and II) and the type of 'additions' defined for the area frame."

Hence, the time t target population is the Wave 1 universe plus "additions." Additions for the area frame sample were civilian, noninstitutionalized United States adults who joined this group by birth, by entering the United States, or by leaving the military or an institution. I expect that the author does not intend to include all such additions in the target population since the Wave 1 universe does not include all civilian, noninstitutionalized United States adults. Maybe only those additions that simultaneously enter the universe and enter a household containing a member of the Wave 1 universe are intended to belong to the target population. In any case, the field procedures did not provide adequate coverage of additional target population members because only adults (aged $\geq$ 16) in the Wave 1 sample were followed when they moved to new addresses during 1979, as was true for the area frame sample.

Two cross-sectional time t estimators of the population total, X_t, were presented. I found it instructive to reformulate these estimators as follows:

$$1.\quad \hat{X}^L_{t,M} = \sum_{k=1}^{N^L_t} W^L_{t,M}(k)\, X_t(k), \qquad (4)$$

where

$$W^L_{t,M}(k) = \begin{cases} \sum_{i\varepsilon S^I_o} [(\beta_k + U_k)\,\pi_i]^{-1} & \text{for the k-th household in the time t sample containing a member from } S^I_o, \\ 0 & \text{for households in the time t sample containing no members from } S^I_o, \end{cases}$$

S^I_o = {Wave 1 list frame (Group I) sample persons},

β_k = Number of Wave 1 list frame (Group I) persons in the k-th time t household,

$$U_k = \begin{cases} 1 & \text{if the k-th time t household contains any additions (Group III persons),} \\ 0 & \text{otherwise, and} \end{cases}$$

π_i = Selection probability for the i-th Wave 1 list frame (Group I) sample person.

$$2.\quad \hat{X}^L_{t,F} = \sum_{k=1}^{N^L_t} W^L_{t,F}(k)\, X_t(k), \qquad (5)$$

where

$$W^L_{t,F}(k) = \begin{cases} \sum_{j\varepsilon S_o} (S_{kj}/S_{kc})\, W_{oj} & \text{for the k-th household in the time t sample,} \\ 0 & \text{for households not in the time t sample,} \end{cases}$$

S_o = {Wave 1 (time t_o) sample households},

S_{kj} = Number of members (Groups I and II) of the j-th Wave 1 sample household who belong to the k-th time t household,

S_{kc} = (Total number of members of Wave 1 sample households who belong to the k-th time t household) + (Number of additions (Group III persons) in the k-th time t household), and

W_{oj} = Unbiased multiplicity-adjusted weight for the j-th Wave 1 household.

The paper notes that the estimators (4) and (5) do not provide unbiased estimates of the population total, X_t. Part of the problem may be that additional (Group III) sample members explicitly enter the weight computations. Since households in the sample must, by definition, contain at least one Group I or II sample member, Group III persons need not explicitly enter the weight computations.

It should be noted that the alternative weights $W^L_{t,M}$ and $W^L_{t,F}$ do not give positive weights to identically the same households. Time t households that contain Group II people, but no Group I people, are given a weight of zero by $W^L_{t,M}$, whereas $W^L_{t,F}$ is positive for these households.

Consideration should be given to defining the person-level target population as simply the original list frame (Group I) persons. Weights similar in definition to (2) and (3) can then be defined that provide unbiased estimates of population totals for this target population. These weights would be essentially the same as the initial family weights used for defining longitudinal family weights for the state Medicaid household survey component of the NMCUES [See Whitmore et al (1982b)].

3. Person-Level Longitudinal Weights for the SIPP by Judkins, et al

Early in the Judkins paper, it is stated that when an interview is missing for a wave and is bracketed by good interviews, imputation will probably be used for the missing wave. Why not use a longer reference period for the next wave interview and collect the data directly, as was done in the NMCUES?

The SIPP universe at any fixed point in time is defined as the persons aged 15 or older who are members of the civilian, noninstitutional United States population, as well as members of the military living on bases with family or living off bases. Dynamic longitudinal features of this universe are:

1. "Additions" - Individuals who were not members of the Wave 1 Universe but became members of the SIPP universe during the panel's 2 2/3 year reference period.
2. "Exits" - Individuals who left the SIPP universe during the 2 2/3 year reference period due to death, moving out of the United States, or going into the military or an institution.

The Wave 1 interview should probe for the occurrence of such events during the Wave 1 reference period. As was true for the ISDP, only Wave 1 sample members are followed to new addresses when they move, and current SIPP survey procedures do not provide adequate coverage of the "additional" target population members. Methods for improving coverage of the "additional" target population members will be discussed later in this section.

The Judkins paper indicates that the ideal annual longitudinal universe is the union of 12 monthly universes. Either this universe or the union of 366 daily universes should be the target population. The problem of analysis of annual statistics when some population members are survey-eligible for less than the full year is noted as one difficulty with this target population definition. I believe that methods exist or can be developed to adequately address this problem. For example, estimation of an annual mean can be based upon the following statistics:

$Y_a(i)$ = Annual income of the i-th sample member while survey-eligible,

$P_a(i)$ = Proportion of the days in the year that the i-th sample member was survey-eligible, and

$W(i)$ = Longitudinal analysis weight for the i-th sample member.

The population totals for Y_a and P_a would be estimated unbiasedly as follows:

$$\hat{N}(a) = \sum_{i=S} W(i)\, Y_a(i), \text{ and} \quad (6)$$

$$\hat{D}(a) = \sum_{i=S} W(i)\, P_a(i). \quad (7)$$

These estimators would have the following interpretation:

$\hat{N}(a)$ = Unbiased estimate of total annual personal income for the target population, and

$\hat{D}(a)$ = Unbiased estimate of the average daily number of members in the target population.

Hence, the ratio estimator,

$$\hat{R}(a) = \hat{N}(c) \,/\, \hat{D}(a), \quad (8)$$

would provide a consistent estimate of the average annual personal income.

Estimation of the population distribution of annual statistics, such as total annual personal income, is somewhat more difficult. The income of a sample member who was survey-eligible only part of the year requires special treatment. The NMCUES defined a time-adjusted income defined for each sample member as

$$TAY(i) = Y_a(i) \,/\, P_a(i), \quad (9)$$

and produced the distribution of these time-adjusted values. Another possibility is to produce separate distributions of annual income for individuals who were survey-eligible for 12 months, 11 months, 10 months, etc. A third possibility might be to simply estimate the annual average monthly income based upon all sample members who were survey-eligible for one month or more, instead of the average annual income.

Four longitudinal weighting procedures are discussed in Judkin's paper. The first procedre defines a longitudinal weight applicable for all longitudinal analyses of an individual's data, irrespective of the analysis time period. A weight of this type is definitely needed for each sample member to facilitate all types of longitudinal analyses. This first procedure gives zero-valued weights to all "associated" sample members. These data are collected mainly to enable family and household analyses. The other procedures attempt to make greater use of the data for "associated" sample members by giving some of them positive weights for particular analysis time periods. Since these "associated" sample members had a chance of inclusion in the Wave 1 sample and were not selected, the bias and variance reduction properties of these procedures would have to be investigated carefully before these procedures could be recommended. Empirical studies based upon the longitudinal data collected by the ISDP, NMCUES, and/or National Medical Care Expenditure Survey (NMCES) could provide the basis for resolving this issue.

A weighting procedure similar to the first procedure in Judkin's paper can provide improved coverage of the target population with some modification of SIPP field procedures. The changes in field procedure would be the following:

1. Each "additional" sample member becomes a "key addition" (i.e., to be followed to the end of the 2 2/3 year panel and receive positive longtudinal weights) if the first household that the person belongs to after entering the universe is a sample household.
2. The Wave 1 sample housing units (and the half-open intervals between sample housing units and next listed housing units) are to be monitored throughout the 2 2/3 year panel for entry of "additional" universe members. If such "additional" people move into one of these housing units and establish their own independent household as their first household after re-entry into the universe, they are also "key additions."

Using this data collection protocol, all longitudinal weights can be based upon selection probabilities for Wave 1 sample households as follows:

1. For each member of a Wave 1 sample household, the longitudinal weight is the reciprocal of the selection probability for that household.
2. Every "key additional" sample member can be linked uniquely to either a Wave 1 sample household or a Time t (time of entry into the universe) sample household. Hence, the longitudinal weight for such a person is either the reciprocal of the selection probability for the uniquely linked Wave 1 household or the Time t cross-sectional weight of the uniquely linked sample household.
3. All "associated" sample members and other "additional" sample members get a weight of zero because they could have been selected into the sample, but were not.

The person-level longitudinal weight adjustment procedures discussed in Judkin's paper seem reasonable. I would, however, only recommend the cross-sectional consistency adjustments to monthly totals and equalization within marriage groups if the adjustments to the post-stratified weights were minor.

4. Household- and Family-Level Longitudinal Weights for the SIPP by Ernst, et al

Ernst presents four longitudinal household definitions for consideration. Preference is indicated for a "Shared Experiences" definition. What is the justification for choosing this definition? More consideration should be given to the question: "What longitudinal household or family definitions are most useful for addressing analysis issues?"

Ernst suggests that longitudinal families not be identified as such but rather that longitudinal households be classified as family and non-family households. The desirability of this approach is questionable. Families that exist either long-term or short-term as multi-family households are potentially important for family-level analyses. Based upon the NMCES and NMCUES experience, it is not especially difficult to divide households into family reporting units for data collection.

Consideration should be given to identifying the properties that one would like all longitudinal households or families to satisfy. Such properties might include the following:

1. Since cross-sectional families are well-defined at any fixed point in time, it may be desirable for the longitudinal families in existence at any fixed point in time to be identical to the cross-sectional families in existence at that same point in time.
2. It may be desirable for changes in household composition that strongly affect family income or program participation to trigger the beginning and ending of SIPP longitudinal families.

Some questions like "What longitudinal family definition is most useful for assessing the effect of divorce on family income?" should be addressed in detail before adopting a SIPP longitudinal family definition. In fact, consideration of how to best address analysis issues may suggest that multiple longitudinal family definitions are needed to satisfy multiple analysis objectives.

The Ernst paper presents five longitudinal household weighting procedures. Each weighting procedure is based on cross-sectional household weights that are equivalent to Huang's "fair share" weight. This appears to be the proper basis for longitudinal household weights.

The need for data for time periods when the longitudinal family is not in the sample is investigated. The need for this additional retrospective or prospective data depends upon both the family definition and the weighting procedure.

Use of longitudinal weights applicable only to specific time periods is discussed as a means for making use of more of the data collected for specific time periods. As noted in the paper, these procedures also tend to require the greatest amount of data for time periods when the family is not in the sample. The variance/bias tradeoff would have to be carefully investigated for these procedures before they could be recommended. Empirical investigations based upon the ISDP, NMCUES, and/or NMCES databases may be useful in this regard. In any case, it is important to have a longitudinal weight applicable for all time periods to enable longitudinal family analyses of all kinds.

One shortcoming of all family weighting procedures suggested by Ernst is that the families spawned by "additional" sample members all get zero weights. The paper states that the first procedure discussed is the procedure used by the NMCUES. This is not exactly true because the NMCUES traced certain types of "key additional" sample members and assigned positive weights to the families spawned by them. The procedures discussed with regard to the Judkins paper are recommended for identifying and tracing "key additional" people. Given these survey procedures, an unbiased "beginning date" type of longitudinal family weightng procedure is presented in Horvitz and Folsom (1980). Review of this paper is highly recommened to everyone interested in longitudinal surveys.

The weight adjustment procedures discussed in the Ernst paper appear to be appropriate and satisfactory for the most part. However, weight adjustment is discussed as a method for compensating for lack of data for specific time intervals, e.g., prior to the first interview or following the last interview. In order to adjust for this type of nonresponse, the NMCUES used attrition imputation procedures. I feel that attrition imputation is a more satisfactory solution because it can address all data missing due to attrition at once and the resulting database is more amenable to analysis. Finally, I would only recommend the final adjustment of longitudinal family weights to monthly controls if the adjustments were a minor.

REFERENCES

Horvitz, D. G. and R. E. Folsom (1980). Methodological Issues in Medical Care Expenditure Surveys. Proceedings of the Section on Survey Research Methods of the American Statistical Association, pp. 21-29.

Whitmore, R. W., B. G. Cox, and R. E. Folsom (1982a). Family Unit Weighting Methodology for the National Household Survey Component of the National Medical Care Utilization and Expenditure Surveys. RTI/1898/06-03F.

Whitmore, R. W., B. G. Cox, and R. E. Folsom (1982b). Family Unit Weighting Methodology for the State Medicaid Household Survey Components of the National Medical Care Utilization and Expenditure Survey. RTI/1898/06-04F.

Wave	Rotation Group	Interview Month	Reference Period							
					1st Quarter 1979			2nd Quarter 1979		
1	1	Feb '79	Nov	Dec	Jan					
	2	Mar '79		Dec	Jan	Feb				
	3	Apr '79			Jan	Feb	Mar			
2	1	May '79				Feb	Mar	Apr		
	2	Jun '79					Mar	Apr	May	
	3	Jul '79						Apr	May	Jun

Figure 1: 1979 ISDP Rotation Group Interview Months and Reference Periods (First Two Waves)

SURVEY OF INCOME

AND

PROGRAM PARTICIPATION:

SESSION IV

This section is comprised of five papers presented in this session which was sponsored by the Section on Survey Research Methods.

MONTH-TO-MONTH RECIPIENCY TURNOVER IN THE ISDP

Jeffrey C. Moore and Daniel Kasprzyk, U.S. Bureau of the Census

The major impetus to the development of the Survey of Income and Program Participation (SIPP) was the need for more detailed and better quality income data than were available through current survey programs--most notably, the March income supplement to the Current Population Survey (CPS) (David, 1983; Ycas and Lininger, 1981). The SIPP itself has only been in the field since October of 1983, so there are not yet sufficient data for a thorough assessment of its performance. However, the precursor to the SIPP, the Income Survey Development Program (ISDP), is an available and underutilized data source offering a wealth of information to researchers with interests in a wide range of SIPP-related issues.

Background

This paper uses the 1979 Panel of the ISDP to examine a particular data quality problem concerning month-to-month turnover in the receipt of various income types. The basic question, first raised by Czajka (1982), is as follows: given six monthly observations over two consecutive survey waves (each of which covers retrospectively a 3-month period), what is the pattern of recipiency turnover in the resulting five pairs of months? Czajka's interpretation of tables prepared for another purpose by Lepkowski and Kalton (1981) was that in survey waves 1 and 2 of the 1979 panel there was "a pronounced tendency for reported program turnover to occur between waves more often than within waves--i.e., between months three and four rather than the four other pairs of months" (p. 93). Moore (1983), however, in a quantitative analysis of the Lepkowski and Kalton tables, failed to find the effect suggested by Czajka.1/

This discrepancy between the two investigations is attributable to differing interpretations of one of the response indicators in the tables--specifically, whether a particular code indicated "no data" (i.e., a case which could not be matched across the two waves) or "no receipt." Notwithstanding this confusion, two additional factors argued strongly for a more careful examination of the issue. First was the issue of completeness. For their work, Lepkowski and Kalton linked only the first two waves of the 1979 panel, leaving untouched waves 3, 4, and 5. A second shortcoming had to do with the quality of the linking operation itself. Lepkowski and Kalton had at their disposal only an early version of the ISDP data file, which contained numerous errors in the person identifier code crucial to the linking of survey records across waves.2/

Subsequent work carried out by Mathematica Policy Research, Inc., apparently corrected the problems with the person identifiers, resulting in the creation of a linked data file which had substantially more matches than the earlier file produced by the Michigan group. In addition, all five relevant waves of the 1979 Panel were included in the linking operation. The remainder of this paper analyzes and discusses tabulations derived from the later "definitive" edition of the 1979 ISDP data file to address more conclusively the issue of within-wave versus between-wave month-to-month income recipiency turnover.

Method and Results

The income types selected for analysis here were identical to the set used in the original Lepkowski and Kalton paper: the two major earned income categories (wage or salary income; self-employment or farm income), and 15 additional sources including all of the major government transfer programs (e.g., Social Security; Supplemental Security Income; unemployment compensation; veterans benefits; Aid to Families with Dependent Children (AFDC); food stamps; etc.). For these major programs, each respondent in two consecutive waves of the ISDP has six monthly observations; we use the term "month-pair" to refer to each pair of successive months. Thus, each set of linked waves includes five month-pairs, which can be designated as 1->2 and 2->3 (within survey wave n), 3->4 (the last month of wave n and the first first month of wave n+1), and 4->5 and 5->6 (within wave n+1). For each income type in each month-pair, a turnover rate ($p_{i(i+1)}$) was calculated as the number of adult sample persons3/ who changed recipiency status with regard to income source X (i.e., who received income of type X in the first month of the pair but not in the second, or vice versa) divided by the total number of adult sample persons. The between-wave rate, p_{34}, was then compared to the average of the within-wave rates, $\overline{p} = 1/4 (p_{12} + p_{23} + p_{45} + p_{56})$. The difference between these two values, $p_{diff} = p_{34} - \overline{p}$, comprises the major variable of interest for this paper.

Table 1 summarizes the results of a simple test of significance4/ carried out on each p_{diff} for the 17 income types across all sets of linked survey waves5/. The message of Table 1 is unmistakeable. There is a strong and consistent tendency toward greater turnover in recipiency between survey waves than between months within a wave. Of the 85 p_{diff} observations in Table 1, 78 are positive (i.e., $p_{34} > \overline{p}$). Sixty-nine of the differences are significantly positive, 51 are significant at the $p<.01$ level or beyond. In contrast, only one difference is significant in the opposite direction.

Almost as obvious as the general trend in Table 1 are its two apparent exceptions. Six of the seven negative difference scores (including the only significantly negative value) are concentrated in two closely related income sources--educational benefits and Basic

(Presented at the ASA Meetings, Philadelphia, PA, August 1984)

Educational Opportunity Grants (BEOG). The only explanation we have for these outliers follows from the fact that they involve one-time payments at the beginning of school terms. Thus, their receipt may be more easily "dateable" than other income sources, and the single payment means that accurate reporting can never produce more between-wave than within-wave turnover. Aside from these relatively weak exceptions, however, it is clear that the great majority of income sources display an exaggerated turnover rate between survey waves. The important question then becomes: Why is this the case?

Discussion

Although it is perhaps the most commonly assumed explanation, response error is by no means the only possible source of the effects observed in this paper, nor is it necessarily the most likely source. In this final section, we briefly examine four potential contributors to greater between-wave than within-wave recipiency turnover: real underlying trends, edit and imputation procedures, person mismatches in linking data from successive survey waves, and response error.

Real underlying trends: Since this investigation is without the benefit of external validating information, we cannot demonstrate conclusively that the observed results indicate "error" as opposed to reflecting accurately real underlying trends in the events being measured. Two facts, however, render the latter hypothesis untenable: 1) a change in economic conditions or eligibility rules could produce an increase in recipiency turnover at a particular point in time, but it is difficult to imagine this happening periodically for a wide range of income types over an extended period of time; 2) the staggered interviewing schedule for the 1979 ISDP Panel (see Ycas and Lininger, 1981) further reduces this likelihood, since each calendar month over the life of the panel served as the first reference month of a wave for one set of respondents, the second reference month for another set, and the third month for a third set. In other words, each reference month in a survey wave combines data from three calendar months, so that any real change effects are present only in diluted form in three reference months.

Edit and imputation procedures: Three processing procedures possibly contributed to greater recipiency turnover between waves than within waves: reformatting edits to simplify and make consistent various data fields, imputation for person nonresponse, and imputation for item nonresponse.

The only known problem with the reformatting edits is that they were carried out independently for each wave; incorrect resolutions in the name of consistency thus may have artificially reduced turnover within waves, while reporting inconsistencies between waves were ignored. Another edit decision which may have contributed to the phenomenon of less turnover within waves than between waves was the following: if at least one "yes" was reported for an income type, and/or if at least one monthly amount was a valid nonzero amount, then any blank monthly recipiency indicators were set to "yes" and any blank monthly amounts were imputed using the average of the amounts reported in other months. The obvious effect of such a procedure is to reduce the apparent amount of change within a wave. Unfortunately, these edits were not identified on the data file. As a result, the extent to which they affected the results presented here is not known, although their combined impact is likely to be small.

Another possible contributor to the observed effect is the treatment of person noninterviews within interviewed households. Because there were, in fact, few such cases (only 298 in Wave 1), an imputation procedure was developed to substitute complete person records for the otherwise missing data. The procedure used reported demographic data as matching variables in a hot-deck assignment. Since each wave's data were processed independently, it is highly unlikely that an individual who was a nonrespondent in each of two consecutive waves would receive the same imputation donor for both waves. Consequently, some spurious wave-to-wave change could occur solely as an artifact of the independent processing.

The same argument applies to the case of item nonresponse within a person's record. The presence of valid data in one wave and the absence of valid data in the next (or *vice versa*) suggests possible problems for between-wave analyses because the ISDP imputation system did not take previous (subsequent) reporting patterns into account. In addition, if a respondent did not provide information for a specific item on two successive waves of interviewing, it is likely that different imputation donors provided the missing data in each wave.

Mismatches: Technically, of course, although respondents do report month-to-month turnover within a survey wave, it is incorrect to refer to respondents' "reports" of between-wave turnover. These events are created by the computerized process which links together the data for specific individuals across survey waves. To the extent that people are incorrectly linked, a certain amount of artifactual turnover may appear in the month-pair which connects the two waves. Preliminary simulation work suggests that mismatching need not be extensive to produce within-wave versus between-wave differences of the magnitudes observed in Table 1. In fact, for most of the income types in this paper, a mismatch rate of 3 percent or less would produce an apparent increase in turnover quite comparable to the observed increase from within-wave month-pairs to between-wave pairs.

It is impossible after the fact to determine the impact of person mismatches on the estimates of between-wave turnover in the 1979 panel. Returning to the discrepancy between the early Lepkowski and Kalton data and the subsequent refined file, one intriguing possibility is that although the former produced fewer matches than the latter, the matches that were completed may have been relatively error-free. If this were the case--that is, if the Michigan group somehow skimmed off the definite matches--then the appearance of heightened between-wave turnover in the later data file may simply reflect increased match errors. Clearly, evaluating the impact of match errors in turnover estimates from the SIPP will require maintaining data on the quality of the match for each person, perhaps in the form of a scale showing the number of variables which were identical across the linked waves.

Response error: Perhaps the most common explanation for the effects observed in this paper involves some form of recall bias. This was certainly Czajka's (1982) assumption. Presumably, a gestalt-like process operates in response to imperfect recall, leading respondents to report receipt for the entire 3-month period of a single wave as having been more stable than it really was. Such a process would work in two ways to produce more reports of between-wave than within-wave turnover: first, by reducing the number of within-wave turnover episodes (see Example 1); and second, by shifting the occurrence of turnover episodes to the between-wave period (Example 2).

	wave n			wave n+1		
Month:	1	2	3	4	5	6
Example 1						
actual receipt:	yes	no	yes	no	yes	no
reported receipt:	yes	yes	yes	no	no	no
Example 2						
actual receipt:	yes	yes	yes	yes	no	no
reported receipt:	yes	yes	yes	no	no	no

Although it is impossible with the available data to evaluate these notions directly, other research has demonstrated effects which appear to be related to the processes hypothesized to be at work here. Goudreau, Oberheu, and Vaughan (1984) report two results of interest from a survey of known AFDC recipients. First, those who failed to report receipt were likely to have received AFDC income for only part of the reference period of the survey. And second, the most common error in reporting income amounts was the tendency to report "the most recent payment for all three months of the reference period when payments actually varied" (p. 184).

A second, related response error possibility can be examined using the present data. According to this explanation, misreports of the type described above, while perhaps representing a general human tendency, are even more likely to occur when the respondent and the subject of the report are not the same person, and especially when different respondents provide the data for two consecutive survey waves. Table 2 summarizes the data regarding the role of proxy response in general, and changing respondents specifically, on elevated between-wave turnover. The results do not present a simple picture, but there is no evidence that self-response in consecutive waves erases the general effect observed in this paper. Note that with only one exception, all differences in column (c) are positive; that is, between-wave turnover is consistently greater than within-wave turnover even when attention is restricted to the constant self-response group.

Nor, in fact, is there consistent support for the weaker argument that self-response might at least reduce between-wave/within-wave turnover discrepancies. As shown in columns (j) and (m), the weight of the evidence is in the opposite direction. Only for the two earned income categories does proxy involvement strongly and consistently produce greater differences as compared to constant self-response.

Why the two general income types produce such disparate results is not clear. A plausible partial explanation--at least for the both-self/mixed-self-and-proxy comparison--is that a true change in recipiency for earned income also changes a person's availability for interview. For example, those who are not employed may be more readily available to be interviewed for self than those who are employed. Receipt of unearned income, on the other hand, is not associated with with the likelihood of finding a person at home; thus, recipiency turnover for unearned income is not associated with a corresponding change in response status.

Conclusion

This paper has demonstrated the existence of some data quality problems in the 1979 Panel of the ISDP, at least when data are examined from more than one survey wave at a time. We have as yet no definitive explanation for these problems, but only a list of possible causes: edit, imputation, and processing procedures; matching difficulties; and response errors. It is likely, of course, that all contributed to the observed effects.

Although modelled in many ways on the 1979 Panel, the SIPP has adopted several modifications which may reduce the problem of heightened turnover in income recipiency between survey waves. First, the SIPP questionnaire includes procedures by which information brought forward from the previous interview can be verified and corrected, if necessary, at the time of interview. The identification and correction of incorrect information was not systematically addressed in the ISDP. Second, the SIPP exercises much tighter control on the sample than did the ISDP, through an improved control numbering system, and improved check-in procedures in Census Regional Offices. These new procedures should help keep mismatches to a minimum in linking consecutive survey waves.

In the future, as SIPP data become available we will monitor them closely for evidence of the type of problem we have demonstrated here. In addition, we will seek to ensure that data which might help pinpoint the cause of the problem (for example, match certainty indicators and edit and imputation flags) are systematically gathered and maintained. We are also planning a more active program of investigation--a record check study matching selected SIPP income receipt and amount data with existing administrative records. Such a study will contribute greatly to our understanding of the quality of SIPP responses, and will provide valuable direction to the development of any ameliorative actions to improve the quality of the SIPP.

Technical Note on Significance Testing Procedures: The following assumptions guided procedures for testing the significance of the between-wave versus within-wave difference in turnover rates:

Suppose five observations have common variance σ^2 and common correlation ρ. Then

the variance of the average of four

$$= \frac{4\sigma^2 + 12\rho\sigma^2}{16} = \frac{\sigma^2}{4}(1+3\rho)$$

and

the variance of the average of four minus the fifth

$$= \sigma^2 + \sigma^2/4\ (1+3\rho)\ 2\rho\sigma^2$$

$$= (5/4)\sigma^2(1-\rho).$$

In this illustrative example, the effect of positive covariance among the estimates is to reduce the variance below the sum of the variances of the two components. For the tests in Table 1, the variance of the difference was estimated by

$$Var_{diff} = 1/N\ [1/16\ (p_{12}(1-p_{12}) + p_{23}(1-p_{23}) + p_{45}(1-p_{45}) + p_{56}(1-p_{56})) + p_{34}(1-p_{34})]$$

where N = the number of adult sample persons in the two consecutive waves and $p_{i(i+1)}$ = the turnover rate for month-pair i and i+1

which ignores all covariances, and thus is likely to be conservative as compared to the illustrative example.

FOOTNOTES

1/ In fact, if the analysis indicated any consistent tendency, it was quite the opposite of that proposed by Czajka--less turnover in the month-pair which linked the two survey waves than in those within a single wave.

2/ Some suggestive evidence on the extent of this problem can be seen in the fact that about 20 percent of the entries in the Lepkowski and Kalton tables are of the "no match" variety, with data available for only one of the two waves. In fact, it was the frequency of this outcome which led Czajka to believe that the supposedly "no match" cases were actually "no receipt," since the code "occurs too often to reflect simply a failure to match records between waves" (Czajka, personal communication, 1983).

3/ Excluded from the tallies are the special subsamples of persons selected from lists of program participants, and persons who were not adult household members during both of the consecutive survey waves. Sample weights were not used for the tallies, and all analyses used the unweighted survey data.

4/ See the Technical Note regarding the procedures for significance testing.

5/ An explanation is in order regarding the last column of Table 1. In the design of the 1979 Panel, a randomly selected one-third of the sample was not administered a wave 4 interview, but skipped directly from wave 3 to wave 5. Thus, the first two sets of linked survey waves--1&2 and 2&3--contain the full respondent sample, sets 3&4 and 4&5 contain two-thirds of the sample, and set 3&5 contains one-third of the sample.

REFERENCES

Czajka, J. "Subannual Income Estimation." Paper 7 in M. David (ed.), Technical, Conceptual, and Administrative Lessons of the Income Survey Development Program (ISDP). New York: Social Science Research Council, 1983.

David, M. "Measuring Income and Program Participation." In M. David (ed.), Technical, Conceptual, and Administrative Lessons of the Income Survey Development Program (ISDP). New York: Social Science Research Council, 1983.

David, M. (ed.) Technical, Conceptual, and Administrative Lessons of the Income Survey Development Program (ISDP). New York: Social Science Research Council, 1983.

Goudreau, K., H. Oberheu, and D. Vaughan. "An Assessment of the Quality of Survey Reports of Income from the Aid to Families with Dependent Children (AFDC) Program." Journal of Business and Economic Statistics, 2 (2), 1984, pp. 179-186.

Lepkowski, J. and G. Kalton. "Investigations of Longitudinal Imputation Strategies for the ISDP Research Panels." Survey Research Center, University of Michigan, 1981.

Moore, J. "Program Participation Turnover in Waves I and II of the 1979 Income Survey Development Program (ISDP) Test Panel." Unpublished Bureau of the Census memorandum, July 26, 1983.

Ycas, M. and C. Lininger. "The Income Survey Development Program: Design Features and Initial Findings." Social Security Bulletin, November 1981, pp. 13-19.

TABLE 1: BETWEEN-WAVE (p_{34}) AND AVERAGE WITHIN-WAVE ($\bar{p}$) MONTH-TO-MONTH RECIPIENCY TURNOVER RATES, AND ANALYSIS OF DIFFERENCE OF RATES (p_{diff}), FOR SEVENTEEN INCOME SOURCES IN FIVE SETS OF LINKED SURVEY WAVES (1979 PANEL, ISDP)

	Linked Survey Waves														
	1 & 2			2 & 3			3 & 4			4 & 5			3 & 5[1]		
Monthly Turnover Rate:	p_{34}	$\bar{p}$	p_{diff}[2]	p_{34}	$\bar{p}$	p_{diff}[2]	p_{34}	$\bar{p}$	p_{diff}[2]	p_{34}	$\bar{p}$	p_{diff}[2]	p_{34}	$\bar{p}$	p_{diff}[2]
<u>Income Type</u>:															
Wage and salary	8.98%	2.99	**	10.10	3.41	**	10.22	3.58	**	7.51	3.01	**	9.03	1.51	**
Self employment and farm	2.99	0.22	**	3.08	0.29	**	2.85	0.26	**	1.99	0.34	**	3.18	0.32	**
Social Security	1.21	0.13	**	1.05	0.11	**	0.98	0.13	**	0.61	0.15	**	1.18	0.08	**
Federal SSI	0.46	0.04	**	0.36	0.04	**	0.20	0.04	*	0.16	0.03	*	0.26	0.02	*
Unemployment compensation	1.12	0.67	**	0.99	0.56	**	0.89	0.59	*	0.97	0.73	*	0.77	0.58	
Veterans benefits	0.31	0.03	**	0.24	0.03	**	0.16	0.02	*	0.16	0.03	*	0.12	0.02	
Workmans compensation	0.47	0.23	**	0.23	0.15		0.26	0.15		0.14	0.15	-	0.26	0.18	
AFDC	0.33	0.07	**	0.34	0.08	**	0.26	0.08	*	0.34	0.06	**	0.46	0.07	**
Child support	0.37	0.14	**	0.35	0.13	**	0.34	0.14	*	0.25	0.13	*	0.26	0.09	*
Employer or union pension	0.27	0.04	**	0.23	0.04	**	0.27	0.05	**	0.20	0.04	*	0.29	0.02	*
Educational benefits	0.25	0.27	-	0.149	0.153	-	0.20	0.14		0.16	0.17	-	0.07	0.16	-
BEOG	0.48	0.21	**	0.09	0.16	⊖	0.30	0.15	*	0.42	0.35		0.24	0.25	-
Food stamps	1.47	0.48	**	1.23	0.34	**	1.21	0.36	**	0.98	0.35	**	1.13	0.28	**
Rental income	0.56	0.04	**	0.76	0.04	**	0.29	0.07	**	0.44	0.08	**	0.14	0.01	*
Assist. from relatives, friends	0.75	0.37	**	0.63	0.15	**	0.40	0.09	**	0.32	0.13	*	0.53	0.15	*
Lump sum payments	1.38	0.95	**	1.48	1.01	**	1.25	0.80	**	1.00	0.93		1.16	1.08	
Incidental or casual earnings	0.74	0.31	**	0.97	0.38	**	1.13	0.43	**	0.93	0.45	**	0.94	0.45	*
Number of Cases	13,157			12,751			8,568			8,639			4,154		

1/ See text footnote 5.

2/ See the Technical Note for a description of significance testing procedures, results of which are symbolized as follows:

**	$p_{34} > \bar{p}$, $z > 3.3$	($p < .01$)
*	$p_{34} > \bar{p}$, $2.0 < z < 3.3$	($p < .05$)
[blank]	$p_{34} > \bar{p}$, $0 < z < 2.0$	(n.s.)
-	$p_{34} < \bar{p}$, $-2.0 < z < 0$	(n.s.)
⊖	$p_{34} < \bar{p}$, $z < -2.0$	($p < .05$)

TABLE 2: AVERAGE BETWEEN-WAVE (p_{34}) AND WITHIN-WAVE ($\bar{p}$) MONTHLY TURNOVER RATES, AND AVERAGE DIFFERENCE OF RATES (p_{diff}), BY RESPONDENT PATTERN IN CONSECUTIVE WAVES; AND COMPARISON OF DIFFERENCES FOR PROXY SITUATIONS VERSUS CONSISTENT SELF RESPONSE (1979 PANEL, ISDP)

	Average[1/] Monthly Turnover Rates (p_{34} and $\bar{p}$) and Differences[2/] (p_{diff})									Comparison[2/] of p_{diff} for for Proxy Versus Self-Response Situations			
	Respondent Pattern in Consecutive Waves									Pdiff (s-p and p-s) minus Pdiff (s-s)		Pdiff (p-p) minus Pdiff (s-s)	
	Self-Self			Self-Proxy & Proxy-Self			Proxy-Proxy						
Income Type	Average p_{34} (a)	Average $\bar{p}$ (b)	Average p_{diff} (c)	Average p_{34} (d)	Average $\bar{p}$ (e)	Average p_{diff} (f)	Average p_{34} (g)	Average $\bar{p}$ (h)	Average p_{diff} (i)	Average effect (j)	Number of + differences (out of 5)[3/] (k)	Average effect (m)	Number of + differences (out of 5)[3/] (n)
Wage and salary	6.54	2.88	3.66	15.52	3.87	11.66	13.38	3.83	9.54	8.00	5	5.88	5
Self employment and farm	2.53	0.31	2.22	3.57	0.29	3.28	3.18	0.19	3.00	1.06	5	0.78	5
Social Security	1.03	0.12	0.90	1.22	0.12	1.10	0.73	0.09	0.64	0.20	4	-0.27	0
Federal SSI	0.32	0.04	0.29	0.24	0.03	0.21	0.20	0.02	0.17	-0.07	1	-0.11	1
Unemployment compensation	0.93	0.61	0.32	1.20	0.81	0.40	0.75	0.51	0.24	0.08	3	-0.08	2
Veterans benefits	0.20	0.03	0.18	0.21	0.03	0.19	0.13	0.00	0.13	0.01	4	-0.05	1
Workmans compensation	0.29	0.18	0.12	0.26	0.20	0.05	0.19	0.12	0.06	-0.07	2	-0.06	2
AFDC	0.41	0.08	0.33	0.30	0.06	0.24	0.11	0.01	0.09	-0.09	1	-0.25	0
Child support	0.43	0.18	0.25	0.14	0.02	0.12	0.04	0.01	0.03	-0.12	0	-0.21	0
Employer or union pension	0.31	0.03	0.27	0.23	0.01	0.22	0.05	0.04	0.01	-0.04	2	-0.26	0
Educational benefits	0.13	0.15	-0.03	0.31	0.28	0.02	0.19	0.17	0.02	0.05	3	0.04	3
BEOG	0.18	0.14	0.05	0.47	0.47	0.00	0.60	0.35	0.24	-0.05	1	0.19	5
Food stamps	1.37	0.45	0.92	1.02	0.25	0.76	0.73	0.12	0.61	-0.16	1	-0.31	1
Rental income	0.46	0.05	0.41	0.46	0.04	0.42	0.32	0.01	0.31	0.01	3	-0.10	1
Assist. from relatives, friends	0.65	0.22	0.43	0.39	0.15	0.24	0.16	0.03	0.13	-0.18	1	-0.29	0
Lump sum payments	1.41	1.10	0.32	1.02	0.68	0.34	0.83	0.65	0.19	0.02	3	-0.13	1
Incidental or casual earnings	0.93	0.41	0.52	0.91	0.33	0.58	1.00	0.38	0.61	0.06	2	0.09	4

1/Average rates are computed as the sum of the rates derived from each set of consecutive waves divided by five.

2/Minor discrepancies in some differences are due to rounding.

3/Columns (k) and (n) provide evidence on the consistency of proxy involvement effects across the five sets of consecutive waves. Entries indicate the number of times the p_{diff} for the proxy situation exceeds the p_{diff} for the self-self pattern.

The Student Follow-up Investigation of the 1979 Income Survey Development Program

Anthony M. Roman, Diane V. O'Brien, U.S. Bureau of the Census

I. Background

The Income Survey Development Program (ISDP) was a research and development program established in the mid-1970's by the Department of Health, Education and Welfare (HEW) in conjunction with the U.S. Census Bureau to prepare for the upcoming Survey of Income and Program Participation (SIPP). The SIPP is the new survey conducted by the Census Bureau designed to satisfy a wide variety of data needs concerning the economic situation of persons and families living in the United States. Data collection for the first SIPP survey, the 1984 Panel, began October 1983.

The major purposes of the ISDP were the same as the goals set out for the SIPP: to improve current estimates of income and income change; to extend the scope and precision of policy analyses for a wide range of Federal and State tax and social welfare programs, and to broadly assess the economic well-being of the population.[1]

The ISDP conducted four field tests. All were experimental in nature as different concepts, procedures, questionnaires and recall periods were tested. The 1979 Research Panel was the largest and most comprehensive research effort conducted by the ISDP.

The 1979 Panel was a nationwide household survey with a total sample of 11,800 households drawn from 130 Census primary sampling units (PSUs). Of this total, approximately 9300 cases were selected from an area sample and 2500 cases were drawn from list samples. Data collection began in February 1979 and ran through June 1980. One-third of the sample households were interviewed each month during the interview period. Information was obtained on household composition, labor force participation, various sources of money and nonmoney income, taxes, assets and liabilities, and other related topics.

The 1979 Panel included many controlled experiments which tested alternatives for basic survey design. The major tests conducted were: household versus individual questionnaire format; self versus proxy respondent rules; and 3-month versus 6-month respondent recall.

As part of the research effort to test respondent rules, one unresolved issue concerned proxy interviews taken for college students not living at their parents' address. In order to test the validity of information collected for this type of proxy interview, an experiment was conducted during the November and December interviews of the 1979 ISDP Panel. This experiment was called the Student Followup Investigation. This paper discusses the objectives, design, and field procedures used for the investigation, and some preliminary results of this experiment.

II. Purpose

Respondent rules during the 1979 Research Panel were to conduct a personal interview for each adult household member 16 years or older. If a self-response interview could not be obtained, the procedure was to accept a proxy interview from another household member who was knowledgeable about the absent person. In this survey, as in other Census surveys, students were considered as members of their parents' households until they established a permanent residence elsewhere. Therefore, the usual procedure for students living away from home while attending school was to treat them as household members who were temporarily absent and obtain proxy interviews from other members of their parents' household.

The fourth interview questionnaire (Wave 4) used during the 1979 Panel contained a special set of questions concerning post secondary educational enrollment and expenses. Thus, this interview seemed especially appropriate for studying the quality of proxy interviews for students, as compared to the student's self interview.

In order to measure the accuracy of information taken from proxy interviews for students living away from home, the fourth interview was first obtained by proxy at the parents' household, and then by self interview at the student's school residence. This self-response interview is referred to as the student followup interview.

There were two basic purposes for conducting the Student Followup Investigation:

1)To obtain the most complete and accurate information possible for items in the Education Expenses section of the Wave 4 questionnaire (such as school enrollment, tuition, fees, and living expenses), and

2)To determine whether proxy respondents at the sampled address are able to provide reliable information on labor force participation, income, education expenses and enrollment for students living away from home. This experiment would be conducted by comparing the information obtained from the proxy interview taken at the parents' home to the self-response interview taken at the student's school address.

III. Procedures

The fourth interview questionnaire used in November and December was designed to identify students living away from their usual residence while attending school. Only students who were actually staying at their school residence (either a dormitory, fraternity house, apartment, etc.) during the time of the November or December interview were eligible for followup.

Census interviewers were instructed to call their regional office within 24 hours of a household interview in which a proxy interview had been administered for a student who was absent and living away at school. The interviewer would then provide the office with the student's name and school address.

Census regional offices were responsible for the control and assignment of the student followup cases. The rules for assigning the cases were essentially the same as the ISDP rules for movers. If the student's school address was within 50 miles of an ISDP PSU, the office assigned the case for interview. As soon as possible thereafter, an interviewer visited the student for an interview. Regional offices were instructed to always employ a different interviewer for the student's interview in order to eliminate any interviewer bias. Additionally, interviewers were instructed to accept only self-response interviews at the student's school address; no proxy responses from roommates or friends were allowed.

IV. Field Results

The analytic universe for the study was the totality of students in the 4th Wave of the 1979 Panel who usually lived away from home and were attending post secondary schools.[2] There were 443 such students identified. Of these, 117 (26.4 percent) were not eligible for interview since the school residence was more than 50 miles from an ISDP PSU and 54 (12.2 percent) were not eligible because the student was staying at home during the time of the 4th Wave interview.

Of the 272 cases assigned, 202 student followup interviews were obtained yielding a response rate of 74.3 percent. Of the 70 noninterviews, 6 were cases in which the parents refused permission for the interviewer to contact the student.

The major reason for the noninterviews was that many students were not staying at their school address (because of Thanksgiving, Christmas and semester breaks) by the time the interviewer received the followup assignment. Although interviewers were allowed until the first week of December to obtain the followup interviews for students identified in November and until the second week of January for students identified in December, many students remained on some type of break later into December and January. This proved to be an inappropriate time of year for conducting interviews with students at their school address. However, in the case of the 1979 Panel, we overlooked this factor in the survey design in order to conduct the experiment in conjunction with the Education Expenses questions, which were set beforehand for the Wave 4 interview.

A recommendation for future studies involving students interviewed at their school address is to obtain the school address in a previous wave's interview. This would allow interviewers more time to contact the student.

V. Preliminary Findings

A. Data Set Creation

The first task in analyzing these data was the creation of a data set of matched responses from the student followup questionnaire and the proxy questionnaire administered during Wave 4 of the ISDP. During the matching process, 35 students (17.3 percent) could not be matched to the Wave 4 ISDP File. Attempts to reconcile the mismatches were unsuccessful. In all but one instance, the most basic identifiers for these 35 students did not exist on the Wave 4 ISDP File. Due to the time elapsed from the initiation of this followup study to the creation of the analysis data set, it has been extremely difficult to find out why these mismatches have occurred. Future studies should be aware of these problems and prepare for them. Omitting the 35 mismatches resulted in a data set of 167 matched responses. These data are analyzed in this report. In all but two instances, the variables analyzed are direct responses to questions on the ISDP form (i.e., they are not in any way computed). The only exceptions are "usual hours worked per week at all jobs" and "total pay before deductions from all jobs last month". These two variables are computed by summing the response from each reported job.

B. Relationship of Student to Proxy Respondent

The relationship of the student to the respondent serving as his/her proxy can be determined in most cases through their relationships to the household reference person. The reference person is that household member who is stated as owning or renting the residence. Table 1 indicates that in 84.4 percent of the cases, the proxy was a parent of the student. This follows the expected pattern.

C. Wage and Salary Comparisons

The ISDP questionnaire was divided into several sections. One section was designed to identify receipt of income types while other sections obtained amounts. Persons were asked a series of wage and salary questions if they indicated in the recipiency section of the questionnaire that they worked at a job or business. One wage and salary record was created containing responses to the set of wage and salary questions for each job named. Thus, if a student had only one employer, a wage and salary record should have been created with the student's responses while another wage and salary record should have been created with the proxy's responses. The reference period used in the ISDP was the previous 3 months, but the wage and salary records were created on a job basis. Therefore, a reported job could have been held at any time during the 3-month reference period. In examining the 167 matched cases of self and proxy responses, the following breakdown of wages and salaries was observed:

83 had at least one self and one proxy record
53 had neither a self nor a proxy record
27 had a self but no proxy record
4 had a proxy but no self record

If one assumes that the self response is correct, then the proxy failed to identify a job held by the student in 27 cases (24.5 percent). This appears to be rather substantial and indicates a potential source of underreporting of wages and salaries with proxy response. The 4 cases in which a proxy record exists while no self record exists may be interpreted as a potential source of misreporting wages and salaries under proxy response.

In attempting to analyze particular wage and salary questions of interest, several conditions must be kept in mind. While 83 matched cases exist with both a self and proxy wage and salary record, the number of cases available for making comparisons for any particular question may be less. There are two primary reasons for this: 1) one interview may have proceeded in a fashion which asked the question of interest and a response was coded, while the other interview proceeded in a fashion which did not ask the question (i.e., due to the various possible skip patterns within the questionnaire), and 2) even though the question of interest may have been asked during both interviews, one may have resulted in a valued response[3] while the other did not. Valued responses are important in evaluating the quality of data obtained in a survey. They indicate both knowledge by the respondent of the investigated subject matter and willingness to cooperate in the survey.

With this in mind, the percentages of coded responses which were valued (i.e., given that a question was asked, the number of times it resulted in a valued response) are presented in Table 2. It is seen that for several wage and salary questions, it is more likely that a self respondent will give a valued response. This is particularly evident with the "usual hours worked per week" and "hourly rate of pay" variables. In all but one instance, when a valued response was not given, a "don't know" was the recorded response.

Table 2 also presents the mean value of self responses for seven wage and salary variables for three particular categories:

1) the proxy could not identify that the student had a job (i.e., no proxy wage and salary record existed but a self wage and salary record did exist)

2) the self response was valued while the proxy response was not (e.g., the proxy most likely responded "don't know"), and

3) both self and proxy responses were valued.

This table demonstrates that a pattern appears to exist in which proxies best identify jobs at which students earn the most money or work the most hours. The smaller the earnings or hours worked, the more likely the proxy will either not be able to identify the job or not be able to answer detailed questions about the job.

Several points should be noted concerning Table 2. The usual hours worked per week may seem rather high for student jobs. This is due to the reference period for these questions extending back into the summer months. Therefore, summer jobs in which the student may have worked 40 or more hours per week will be included in these summaries. This also explains the decreases in total monthly pay from three months ago to last month. Also, it is impossible to compute total monthly pay by using the usual hours worked per week and regular hourly rate of pay. This is because the values presented in these tables are means and concern the student's primary job. One student's primary job may have been three months ago while another's may have been last month.

The final table, Table 3, presents comparisons of the self and proxy valued responses. It should be noted that the estimated variances used in computing these confidence intervals do not take into account any sample design effects. The reason is that this analysis is considered preliminary and will be used to decide if a more lengthy detailed analysis seems warranted. The net result is that the intervals in Table 3 should be considered conservative, while conclusions of significant differences should be considered liberal. Computation of design effects may add a small degree of accuracy to results from this study, but it should be noted for future studies, that increased emphasis on obtaining responses from all sample students and their proxies would greatly enhance the accuracy of results. Of the seven wage and salary variables analyzed, two showed a significant difference at the .05 level. These were "usual hours worked per week" and "regular hourly rate of pay", both for the student's primary job. In both instances, the proxy gave the larger valued mean response. It is interesting to note that for "usual hours worked per week at all jobs", the mean self and proxy responses are not significantly different. This raises the question of the proxy and student possibly identifying different jobs as being primary.

D. Education Expenditure Comparisons

All 167 matched cases had both a self and a proxy education expenditures record, but 61 of these records were unavailable for this preliminary analysis. This was due to a flaw discovered in the manner in which the Wave 4 ISDP data were processed. Only rekeying of the questionnaires could retreive the data and this was deemed unwarranted at the present time. Therefore, 106 matched records were available for analyzing education expenditures.

Table 2 again presents the percentage of responses which were valued. It is obvious that a valued response is much more likely from a self respondent than a proxy respondent. This seems understandable for all variables except "amount paid by family on tuition and fees" since the other variables involve expenditures most likely handled directly by the student. In every instance that a valued response was not given, "don't know" was the recorded response.

Table 2 displays the mean value of self responses both when the proxy has a valued response and also when the proxy response is "don't know". Three of the four variables considered do not appear to differ substantially between these two categories. Only the "amount paid by family on tuition and fees" exhibits a rather large difference with the mean self response being greater if the proxy has a valued response. This is consistent with the wage and salary results in that the more expensive the tuition, the more the proxy is likely to know about the amount. It may also help explain why so many "don't knows" were given by proxies in response to this question. Perhaps when the amount of tuition is low, the student is more likely to be directly involved in its payment(e.g., the student may pay the tuition from support supplied by the parent).

Table 3 again presents results of comparisons of self and proxy valued responses. Two of the four variables showed a significant difference at the .05 level. They were "academic credit hours taken this term" and "cost of course materials". In both instances, the mean proxy response was larger. The mean proxy response was also larger for "amount paid on tuition and fees" but with a large estimated variance, a statistically significant difference could not be detected.

E. Other comparisons

Two additional areas were investigated in this preliminary analysis. The first was educational assistance. Of eight assistance categories, only two had enough reported cases to analyze questions concerning amounts received. These were: Basic Educational Opportunity Grants (31 cases) and Government Scholarships, Fellowships, Etc.(11 cases). The results of comparisons of self and proxy valued responses are shown in Table 3. No significant differences in mean amount received were found for any of the assistance variables.

The last area investigated was receipt of interest income. Reporting of interest was handled in the ISDP questionnaire in the same manner as wages and salaries. That is, a person was asked a series of questions regarding amounts of interest if they indicated receipt of interest income in the recipiency portion of the questionnaire. For the 167 matched cases, the reported interest was as follows:

104 cases had both a self and proxy report
30 cases had neither a self nor proxy report
27 cases had a self but <u>no</u> proxy report
6 cases had a proxy but <u>no</u> self report

Assuming the self response is correct, the proxy failed to identify that the student would have interest income in 27 cases (20.6 percent). Although this appears to be a large problem, interest income is poorly reported for all people. For example, in the 104 cases in which both the self and proxy respondent reported receipt of interest earned on the student's own accounts, 61.0 percent of the coded self responses were "don't know" while 81.5 percent of the coded proxy responses were "don't know". Considering the question on interest earned on the student's shared accounts, 69.4 percent of the coded self responses and 80.0 percent of the coded proxy responses were "don't know". Obviously, it appears that the quality of interest data for students is suspect regardless of whether a self or proxy interview is conducted.

VI. Conclusions

The aim of this preliminary analysis was to examine the self and proxy student data in order to decide if a more extensive investigation (e.g., effects of accepting proxy responses on overall survey estimates) seemed warranted. Any inferences drawn from these data should keep in mind that the estimated variances did not reflect any sample design effects and that the size of the data set is quite small. Indeed, most comparisons were based on less than 100 observations. Still, this study is unique and although somewhat flawed in administration and implementation, it is possible to make certain general remarks. When valued responses are available from both the self and proxy interviews, the quality of the proxy responses appears to be generally quite good. Substantially more data would be needed to derive better estimates of the difference between self and proxy response and to narrow the confidence intervals around these estimates.

A problem that does appear to exist is in obtaining a valued proxy response. Quite often, a proxy cannot identify a particular source of student income (e.g., wages and salaries) and even if they can identify it, they are more likely to respond "don't know" to the particulars about that source. A trend does seem to exist that the larger the income or expense, the better the proxy response becomes. Still, this implies that by using proxy responses, the lower range of income or expense amounts are more likely missed.

Finally, the main issues involved in interviewing students away from home are the impact of accepting proxies on overall survey estimates and the differential costs involved in obtaining self responses. Since no cost data are available from this study, an estimate of the additional amount required in obtaining self responses cannot be computed. It may be possible to make some very general comments about the potential impact of accepting proxies on overall survey estimates. Students living away from home make up less than 3 percent of the overall ISDP sample. With this in mind and the fact that results from this study indicate that proxies are more likely to miss only the smaller expense and income amounts, it may appear unlikely that overall survey estimates will be strongly affected. Still, the limitations of the sample involved in this study must be considered in any statement of results. For instance, students living more than 50 miles from an ISDP PSU were omitted from consideration. Also, problems were encountered in matching students to proxies and in losing some survey data due to a processing flaw. The effect that these students could have had on results from this study is unknown. In concluding, further detailed investigation of this particular data set is not recommended due to the limitations in the size and composition of the sample. Future study may lead to stronger results but based upon this preliminary investigation, it is recommended that while the self-proxy student issue should not be forgotten, it should not occupy a high place on the SIPP research agenda.

FOOTNOTES

1 Research Triangle Institute. 1983. The 1979 ISDP Research Panle Documentation. National Technical Information Service, Washington, D.C.

2 Since Wave 4 of the 1979 Panel was administered over a two month period, only two-thirds of the 11,800 household sample was interviewed, making the Wave 4 sample size approximately 8,100 households.

3 Throughout this report, the term valued response is used to imply any response with a legitimate value for the question asked. Valued responses do not include refusals, don't knows, or responses whose value is considered out of range or in some other manner erroneous.

Table 1: Student/Proxy Relationships to Reference Person

student:	Proxy: ref. person	spouse of ref. pers.	child of ref. pers.	other rel. of ref. per.	unknown
ref. pers.	–	0	0	2	1
child of ref. pers.	60	81	6	0	11
other rel. of ref. pers.	3	1	2	0	0
Total	63	82	8	2	12

Table 2: Results for Wage & Salary and Education Expenditure Variables

	% of coded cases which were valued:[1]		Mean value of self responses when:[2]		
Variable	self response	proxy response	proxy could not identify that student had a job	proxy did not give a valued response	both self and proxy had valued responses
Wage and Salary					
Usual hours worked per week at primary job	98.7% (n=76)	76.4% (n=55)	22.30 hrs (n=20)	21.11 hrs (n=38)	35.60 hrs (n=37)
regular hourly rate of pay at primary job	100.0 (n=66)	73.3 (n=75)	$3.17/hr (n=16)	$3.46/hr (n=18)	$3.39/hr (n=48)
total pay before deductions from primary job 3 months ago	100.0 (n=60)	100.0 (n=51)	$111.81 (n=16)	$246.48 (n=27)	$378.42 (n=33)
total pay before deductions from primary job 2 months ago	100.0 (n=61)	100.0 (n=50)	$32.19 (n=16)	$97.82 (n=28)	$138.33 (n=33)
total pay before deductions from primary job last month	100.0 (n=61)	98.0 (n=50)	$37.25 (n=16)	$74.83 (n=29)	$100.00 (n=32)
usual hours worked per week at all jobs	–	–	25.76 hrs (n=21)	24.75 hrs (n=37)	41.37 hrs (n=38)
total pay before deductions from all jobs last month	–	–	$94.47 (n=17)	$131.10 (n=63)	$114.53 (n=51)
Education Expenditures					
academic credit hours taken this term	98.0 (n=101)	50.0 (n=102)	– –	14.88 hrs (n=49)	14.80 hrs (n=50)
amount paid by family on tuition and fees this term	95.0 (n=100)	74.5 (n=102)	– –	$469.25 (n=24)	$1004.10 (n=71)
cost of course materials this term	100.0 (n=100)	46.1 (n=102)	– –	$105.59 (n=54)	$98.75 (n=44)
amount of monthly rent and utilities	100.0 (n=30)	66.7 (n=33)	– –	$148.67 (n=9)	$131.14 (n=21)

1 In these two columns, the numbers in parentheses are amounts of coded cases.
2 In these three columns, the numbers in parentheses are amounts of valued responses.

Table 3: Summary of Comparisons

Variable	Mean Self Response	Mean Proxy Response	Difference	95% Confidence Limits [1] Lower	Upper
Wage and Salary					
usual hours worked per week at primary job	35.60 hrs	40.57 hrs	-4.97 hrs*	-9.03 hrs	-0.91 hrs
regular hourly rate of pay at primary job	$3.39/hr	$3.54/hr	-$.15/hr.*	-$.29/hr	-$.01/hr
total pay before deductions from primary job 3 months ago	$378.42	$336.09	$42.33	-$56.80	$141.46
total pay before deductions from primary job 2 months ago	$138.33	$121.52	$16.81	-$30.79	$64.41
total pay before deductions from primary job last month	$100.00	$106.56	-$ 6.56	-$22.56	$ 9.44
usual hours worked per week at all jobs	41.37 hrs	40.55 hrs	0.82 hrs	-3.55 hrs	5.19 hrs
total pay before deductions from all jobs last month	$114.53	$114.38	$.15	-$21.65	$21.35
Education Expenditures					
academic credit hours taken this term	14.80 hrs	16.00 hrs	-1.20 hrs*	-1.93 hrs	-0.47 hrs
amount paid by family on tuition and fees this term	$1004.10	$1157.63	-$153.53	-$591.87	$284.81
cost of course materials this term	$98.75	$120.84	-$22.09*	-$39.01	-$5.17
amount of monthly rent and utilities	$131.14	$121.38	$9.76	-$16.87	$36.39
Education Assistance					
BEOG assistance received 3 months ago	$299.61	$381.74	-$82.13	-$281.99	$117.33
BEOG assistance received 2 months ago	$194.48	$301.24	-106.76	-$302.28	$ 88.76
BEOG assistance received last month	$ 50.57	$ 14.71	$35.86	-$ 10.11	$ 81.83
Government scholarship assistance received 3 months ago	$ 87.55	$107.73	-$20.18	-$ 63.27	$ 22.91
Government scholarship assistance received 2 months ago	$154.75	$219.75	-$65.00	-$151.52	$ 21.52
Government scholarship assistance received last month	$ 8.00	$ 28.50	-$20.50	-$ 66.87	$ 25.87

* implies difference is significant at the .05 level

1 These limits are based on variance estimates which do not take sample design effects into account.

THE ISDP 1979 RESEARCH PANEL AS A METHODOLOGICAL SURVEY: IMPLICATIONS FOR SUBSTANTIVE ANALYSIS

Richard A. Kulka, Research Triangle Institute

The 1979 Research Panel was the third of three major field tests conducted by the Income Survey Development Program (ISDP) to fulfill its mandate to examine and resolve the technical and operational problems involved in the design and implementation of a survey mechanism that would meet the needs for improved data on income, assets, and program participation for program and policy analysis. As the most proximal and realistic pilot or prototype for the Survey of Income and Program Participation (SIPP), the 1979 Research Panel employed a longitudinal panel design, whereby persons at sample addresses were contacted early in the calendar year and recontacted at regular intervals (usually every three months) and asked about their income and other characteristics for the preceding few months. In this way, a highly detailed record was built up for each person for the entire calendar year according to the schedule presented in Table 1. In addition, since less time was required to update basic information after the initial interview, time was available in later waves of interviewing to ask additional questions or questionnaire modules on: (a) topics of interest that were stable enough not to require updating on each visit; and (b) emerging issues of special interest to particular agencies or programs. As a result of the diligent application of this modular approach, the fully-implemented 1979 Research Panel produced an overwhelming array of data suitable for longitudinal analysis and a corresponding wealth of detailed socioeconomic data on more specialized issues suitable for complex cross-sectional analyses.

Table 1
1979 Research Panel Interview Months and Reference Period

Wave	Rotation Group	Interview Month	Three Month Reference Period
1	1	Feb. 1979	Nov. 1, 1978 - Jan. 31, 1979
	2	Mar. 1979	Dec. 1, 1978 - Feb. 28, 1979
	3	Apr. 1979	Jan. 1, 1979 - Mar. 31, 1979
2	1	May 1979	Feb. 1, 1979 - Apr. 30, 1979
	2	Jun. 1979	Mar. 1, 1979 - May 31, 1979
	3	Jul. 1979	Apr. 1, 1979 - Jun. 30, 1979
3	1	Aug. 1979	May 1, 1979 - Jul. 31, 1979
	2	Sep. 1979	Jun. 1, 1979 - Aug. 31, 1979
	3	Oct. 1979	Jul. 1, 1979 - Sep. 30, 1979
4	1	Nov. 1979	Aug. 1, 1979 - Oct. 31, 1979
	2	Dec. 1979	Sep. 1, 1979 - Nov. 30, 1979
5	3	Jan. 1980	Oct. 1, 1979 - Dec. 31, 1979
	1	Feb. 1980	Nov. 1, 1979 - Jun. 30, 1980
	2	Mar. 1980	Dec. 1, 1979 - Feb. 29, 1980
6	3	Apr. 1980	Calendar Year 1979 (Jan. 1, 1979 - Dec. 31, 1979)
	1	May 1980	
	2	Jun. 1980	

As the largest, most complex, and most "complete" of the ISDP field tests, the 1979 Research Panel is of sufficient design, size, and interest to occupy social and policy analysts for many years to come, and its use for such substantive and policy research has been actively promoted (e.g., David, 1983; Kasprzyk, 1983a). It is important to keep in mind, however, that the fundamental purpose of the ISDP field tests--including the "state-of-the-art" 1979 Research Panel--was methodological: to serve as a flexible vehicle for conducting field experiments and feasibility tests to evaluate the effectiveness of alternative design features and data collection strategies (Ycas and Lininger, 1981). Hence, the 1979 Research Panel is replete with such methodological assessments, and potential users of these data should be aware of their nature for at least two reasons. First, since most of these tests and experiments have not been fully analyzed or evaluated (some not at all), the potential for methodological as well as substantive analysis of these data is very great (cf. David, 1983). Second, since these methodological assessments are an integral part of the total survey design, they bear directly in some cases on the likely quality of the data for substantive analysis, including: (a) potential differences in the nature or quality of data collected under different experimental variations; (b) the confidence one can have in the accuracy of certain data; and (c) the extent to which data collected under different procedures can legitimately be conceptually merged for certain analyses rather than analyzed separately.

OVERVIEW OF TESTS AND EXPERIMENTS

With the latter especially in mind, I have summarized in Table 2 the full range of explicit methodological research embedded in the 1979 Research Panel. Although virtually every aspect of the 1979 Research Panel was subject to methodological scrutiny and evaluation, in particular, five formal, controlled experimental comparisons of alternative design or data collection strategies were systematically incorporated in the survey, and seven other procedures were explicitly included to provide a focused nonexperimental assessment of their feasibility for implementation in the SIPP. For each of these 12 tests and experiments a brief description is provided, along with its basic design, a capsule summary of results, and a note on its possible implications (if any) for substantive analysis.

Controlled Experiments

The first of five formal experiments included in the 1979 Research Panel compared two alternative questionnaire formats for measuring income recipiency, one using a "household screening approach" to determine receipt of various kinds of income and the other a more conventional person-by-person "individual" approach. It was hoped that the former approach would reduce the time needed to administer the questionnaire without a corresponding reduction in data quality. Preliminary analyses by Coder (1980) and Kaluzny (1981) indicated few differences between the two approaches in estimates of income recipiency rates by type, and only a slightly higher incidence of "don't knows" and "refusals" under the household screening approach, but the average savings per household was only about five minutes.

While differences between these two experimental groups are apparently small and undra-

Table 2
The 1979 ISDP Research Panel:
Overview of Tests and Experiments

Test or Experiment	Description	Study Design	Capsule Summary of Results	Implications for Substantive Analysis
CONTROLLED EXPERIMENTS				
1. Alternative Questionnaire Formats	Experimental comparison of a "household screening" versus an "individual screening" approach for measuring income recipiency.	Random assignment to households of alternative forms (one-third to household approach) during first wave of interviewing only.	Preliminary results suggest that household screening format did not offer significant improvement over the individual approach in either time or data quality.	Data on income recipiency gathered during Wave 1 confounded by any potential form effects.
2. Self-Respondent Versus Proxy Respondent Rules	Experimental comparison of rules maximizing self-response with those permitting proxy interviews for absent household members.	Random assignment of households to respondent rule alternatives (one-third to self-respondent) throughout all waves of interviewing.	Slightly higher costs and somewhat better data under self-response rules vs. higher interview rates under proxy treatments.	Data gathered in all interview waves subject to additional variation due to respondent rule treatments.
3. Three-Month Versus Six-Month Recall of Asset Income	Experimental comparison of length of recall period for property income.	Assignment of random half-samples to alternative reporting periods in Waves 1-5.	Little analysis done to date.	Data on amounts of asset or property income potentially subject to variation due to length of recall period.
4. Variable Recall of Monthly Information	"Staggered" interview design providing systematic comparison of monthly information reported using a 1- , 2-, and 3-month reference period.	Equal groups of households randomly assigned to a different month for each wave of interviewing.	Preliminary analyses reveal no consistent evidence of systematic recall effects.	Staggered design complicates calculation of monthly and calendar quarter estimates. May reduce recall bias in estimates of monthly income.
5. Seven-Point Versus Ten-Point Attitude Scale	Experimental comparison of reports of subjective economic well-being as assessed by two alternative attitude scales.	"Split-ballot" technique with a random half of self-respondents responding to each scale in Waves 1-3.	Ten-point scale resulted in broader distribution, reduced positive skew, and increased valid variance.	Measures of subjective economic well-being measures in Waves 1-3 subject to increased variance due to response scale differences.
FEASIBILITY TESTS				
6. Student Follow-up Interviews	Examination of accuracy of information collected for students living away at school during interview period.	Administered Wave 4 questionnaire twice: by proxy at parents' address and in person at school address.	Preliminary analyses suggest quality of proxy response is reasonably good when proxies are able to report information.	Proxy data on college students somewhat incomplete, but not seriously so.
7. Nonfarm Self-Employment Income Test	Tested feasibility of using off-line mail-back surveys for obtaining quarterly estimates of self-employment income.	Respondents self-employed in a business or professional practice during Wave 1 mailed self-response mail questionnaire in first two calendar quarters.	Poor response rates forced abandonment of test after second calendar quarter.	Data on subannual self-employment inadequate for 1979 Research Panel.
8. Staggered-Interview Design	Study to determine whether increased interviewer experiences with the questionnaire and survey resulted in lower error rate.	Interviews conducted quarterly with one-third of the households interviewed during each of the three months comprising a quarter.	Little evidence that monthly interviewing results in substantially improved field performance.	See "4. Variable Recall of Monthly Information."

Table 2 (Continued)

Test or Experiment	Description	Study Design	Capsule Summary of Results	Implications for Substantive Analysis
9. Social Security Number Validation	Test feasibility of obtaining valid SSN's for sample household members.	Two rounds of validation, both including a computer match and manual search of administrative records, and field follow-up on invalid or missing SSN's.	Valid SSN's determined for 95.5 percent of cases included in project.	Survey information capable of being linked with administrative records systems using SSN identifier for high proportion of sample persons.
10. SSI Administrative Data Match	Match of survey and administrative records to validate information common to both sources and enhance survey data base.	Computer and clerical match of records using SSI subsample and SSI administrative tapes.	Final match rate of 99 percent achieved, but no analysis of data quality conducted to date.	None: list frame sample records not included on working files.
11. SSI Domain Match	Match to determine number of persons included in panel through area and BEOG subsamples also in SSI subsample frame.	Computer and clerical match of sample members with valid SSN's to SSI universe file.	Time consuming process resulting in "reasonable" match rate.	None: only area frame cases included in public use files. No need for multiple frame estimation.
12. Mover's Cost Study	Assessment of costs of following sample individuals moving to new address within 50 miles of an ISDP PSU.	Systematic recording by interviewers of their mileage and time spent in discovering, locating, and following up movers.	Mover household follow-up rate of 76 percent and a cost increase of about 8 percent attributable to following movers.	Substantial reduction in nonresponse due to loss of mobile sample members.

matic, however, there is nevertheless little doubt that data on income recipiency gathered during Wave 1 of the 1979 Research Panel is confounded somewhat by any such potential "form effects". By definition the "household screening approach" compelled a higher level of "proxy" rather than self-reporting for income recipiency than under the individual approach (even under standard respondent rules), and because a few questions on income receipt that could not be covered on a household basis (e.g., verification of labor force, retirement, and disability status and Medicare and Medicaid coverage), and all questions concerning amounts of income received, were asked on a person-by-person basis even under the household screening approach, there is a distinct possibility that different types of income data may have been reported differentially under that approach but not when the individual approach was used. Moreover, although these alternative forms were used only during the first wave of interviewing, one cannot rule out the possibility of a longitudinal "forms effect" on these data. For example, serving as the "household respondent" in Wave 1 might result in more accurate reporting of one's own economic data across time than would be the case if each person reported income recipiency for themselves, and, conversely, those initially not serving as household respondents might subsequently be poorer reporters of income recipiency in Wave 2 when exposed to these questions for the first time.

The second experiment involved a controlled comparison of alternative respondent rules. Some households were interviewed using rules that required self-response except in special situations, while others were interviewed under "standard" respondent rules, whereby proxy interviews were accepted for absent persons from other household members when convenient. A number of different analyses have been conducted to date in an effort to study the effects of these proxy respondent rules and self-respondent rules on data quality, noninterview rates, and costs of data collection (e.g., Coder, 1980; Kaluzny, 1981, 1982; Kulka, 1983). In general, while the use of self-response rules results in approximately 20 percent more self-response (85 vs. 65 percent) and 4-6 percent higher interviewing costs than standard respondent rules, results on nonresponse and data quality are mixed. While the proxy treatment had a positive effect on household and person interview rates, self-respondent rules apparently resulted in somewhat better data (as implied, for example, by the greater use of records, lower item nonresponse for certain key items, less rounding, and less variance in non-zero amounts), although some of these effects appeared to be somewhat smaller by Wave 2 (Kaluzny, 1982).

Unlike the "forms" experiment, the self-proxy response experiment continued throughout all waves of data collection, so that any potential influences on data quality or increases in the variance of key variables due to this experimental factor are likely to be found throughout the database. Yet, with few exceptions, the longitudinal implications of these alternative respondent rules have not yet been investigated. Changes in the proportions of proxy respondents or in the characteristics of proxy vs. self-responders under the two conditions may vary over time, thereby confounding somewhat longitudinal analyses of variables especially sensitive to respondent rules. More generally, suppose that a comparison of the 1979 Research Panel to other survey data suggested that the former provided more accurate data relative to an independent source of information. Suppose further that this improvement was directly attributable to the use of a maximum self-response rule (i.e., under regular proxy rules estimates were similar). Without making such an assessment, however, one might assume that <u>in general</u> the ISDP design results in better data and generalize this assumption to the SIPP, an erroneous presumption, of course, unless the SIPP were to employ rules maximizing self-response. Moreover, it is not difficult to see how such a methods "artifact" might similarly influence important relationships among variables of major policy significance.

The third experiment compared property or asset income amounts reported using a three-month reference period with that reported for a six-month recall period. The basic objective of this experiment was to determine if information on asset or property income data collected every six months would be as accurate as that collected quarterly. Results of this experiment would provide evidence on the magnitude of loss with the longer recall period (a critical ingredient in justifying the current four-month recall design for the SIPP), but very little of this analysis has been done to date (cf. Czajka, 1983).

The very reason for conducting this experiment, however, implies increased variation in reported amounts of asset or property income due to differences in length of recall period. Since the preceding three months are reported with an identical recall period by both groups every other wave, the influence is not constant for all months. Thus, substantive analyses of a "common" three-month period may yield different results than that of a similar period where recall is three months longer for one group than the other. Similarly, quarter-to-quarter variation in asset income reporting may be greater within the six-month reporting group than within the three-month subsample. Moreover, since asset income <u>recipiency</u> is reported quarterly, the expected influences would likely be on asset income amounts, but a longer reporting period for "amounts" could also have an indirect adverse effect on reports of recipiency as well. In addition, if a recall effect of either type is present, such effects may either dissipate or increase in magnitude over the life of the panel (through Wave 5).

A fourth experiment, afforded by the use of a "staggered" interview design in which each quarter's interviewing was spread over three months with a variable three-month reference period (see Table 1), provided for a systematic comparison of income and other information reported for several months during the year using a one-, two-, or three-month reference period. Although the staggered design was not adopted for this reason, it provides a "natural" experimental design for the assessment of potential monthly recall bias by length of reporting period for virtually all income types and a wide variety of other variables. To date, however, only preliminary analyses of this natural recall experiment have been conducted (Kaluzny, 1981, 1982; Czajka, 1982), none of which have provided consistent evidence of a systematic recall effect.

Nevertheless, the implications of this monthly interview design for substantive analyses are considerable. On the positive side, the staggered interview procedure provides an ongoing measure of monthly recall bias, and to the extent that such bias exists, the varied recall period tends to minimize its effect (relative to more typical quarterly interviewing) when making comparisons of monthly changes, since income and other monthly data were always collected with the same average length of recall. On the other hand, the staggered approach introduces some substantial problems with regard to missing data and response variance for monthly and quarterly estimates. Point estimates for a given month are made with higher variation, and the staggered approach requires that calendar quarter estimates for two thirds of the sample be derived from data collected in two separate interviews, resulting in greater levels of missing data, linkage problems, and increased month-to-month variation within quarters. For example, recent analyses of data from the 1979 Research Panel indicate a degree of variation in quarterly earnings greater than seems reasonable, and month-to-month changes in income recipiency generally tend to be greater <u>between</u> interviews than in the reference period reported <u>within</u> each interview (David, 1983:11; Moore and Kasprzyk, 1984).

A final area of controlled experimentation involved a comparison of reports by self-respondents of their subjective economic well-being assessed by using either a seven- or ten-point attitude scale. Preliminary analyses of these data (Vaughan and Lancaster, 1980) suggested that the ten-point scale achieved the desired effect of spreading out their distributions and reducing the positive skew associated with the seven-point scale, and additional analyses with a few variables suggested that this was primarily valid variance. In light of these findings, it would appear that these economic well-being measures may indeed vary significantly according to which response scale version is used, and substantive analysis involving these variables must take such variation into account, either by conducting separate analyses with these alternative measures or by directly including in the analysis a variable indicating which of these two experimental versions was used.

<u>Feasibility Tests</u>

In addition to the formal experimental comparison of self-respondent versus proxy respondent rules, two other more specialized respondent tests were carried out. One examined the accuracy of information collected for students living away at school during the interview period by administering the fourth wave questionnaire twice for absent students -- once by proxy at the parents' address and a second time in person at the school address. The basic objective of this study was to evaluate both differences in reporting and the additional burden imposed on field staff when students were followed to their temporary addresses. With regard to the latter, over one-fourth of the students identified lived at a school address outside the sample area, and of those assigned for follow-up, only 74 percent were interviewed, with most nonresponse due to inability to contact respondents at their school addresses. Preliminary analyses of data from the students interviewed indicate that when "amounts" or details are available from both the self and proxy interviews the quality of proxy responses is generally quite good, but proxy respondents are frequently unable to provide a "valued response" at all (cf. Roman and O'Brien, 1984). In general, then, proxy data obtained for college students are clearly somewhat incomplete, but most analyses of data from the 1979 Research Panel should not be greatly influenced by these deficiencies, with the possible exception of those which rely on special subsamples containing a large proportion of college students and focus on variables especially prone to such proxy reporting error.

The second respondent test examined the feasibility of using off-line mail-back surveys for obtaining quarterly estimates of nonfarm self-employment income from respondents owning a business or professional practice. Because of poor response rates, this particular effort to measure subannual self-employment income was abandoned after the second quarter. Although some substantive analyses have been conducted using these data (e.g., Whiteman, 1983), methodological analysis took the form of additional experimentation with alternative procedures in an effort to improve this performance, none of which were very successful. The major implication of this feasibility test for social or policy analysts is that data on subannual self-employment income collected in the 1979 Research Panel are generally regarded as deficient.

The staggered interview design (mentioned earlier), which roughly tripled each interviewer's experience with a form, was itself a feasibility study. In addition to routine quality control interviews, an expanded reinterview program was initiated to determine whether such increased interviewer experience with the questionnaire and with the survey in general results in lower error rates. Research conducted to date provides little support for the proposition that monthly interviewing resulted in substantially improved field performance or data quality. Should such differences exist, however, this "interviewer experience" or "learning curve" factor may combine with potential recall factors to increase variation in these survey data, variation which would not be there if all interviews were conducted the first week of the calendar quarter, for example.

Two other feasibility tests incorporated in the 1979 Research Panel were designed to explore issues related to linkage of survey responses with data in administrative records systems. First, since the Social Security Number (SSN) is the identifier in most general use, a project to determine valid SSN's for sample households was conducted using two rounds of validation, both including a computer match and manual search of Social Security Administration (SSA) administrative records. Through the use of those procedures and exploiting the panel design to obtain corrected SSN's from the field in later interview waves, valid SSN's were determined for 95.5 percent of the cases included in the project, a rate that might be improved with minor modifications in the future (Kaspryzk, 1983b). As a result, should access to administrative records systems be granted, substantive analysis using survey information linked to records data would be possible for a high proportion of persons sampled in the 1979 Research Panel.

Second, two distinct projects were undertaken to examine the feasibility of linking 1979 Research Panel data to benefit records of the Supplemental Security Income (SSI) program. The first involved a match of survey and administrative records using the 1979 Research Panel SSI subsample and SSI administrative tapes in order to validate information common to both sources and enhance the survey database. Overall, 3,950 sample persons in the 1979 Research Panel were matched with the SSI data sets, yielding a final match rate of 99 percent. However, analyses of data quality on this survey-administrative data match have not yet been conducted, and, since these list frame sample cases are not included in the public use microdata files (NTIS, 1983), this project is of no particular concern to substantive data analysts. Similarly, the second linking project, an SSI "domain match", was designed to determine the number of persons included in the panel through the area and Basic Educational Opportunity Grants (BEOG) subsamples who were also in the frame used to select the SSI subsample. Employing a match indicator code algorithm using validated SSN's, in combination with name, race, and date of birth, a reasonable match rate was achieved, albeit over a longer time period than would be required to support multiple frame estimation (Kaspryzk, 1983b). Since only the area frame cases are included in the public use files, however, multiple frame estimation is not required for substantive analyses of these data.

Finally, in an effort to determine the incremental costs of following movers (an integral feature of the survey design for the 1979 Research Panel and the SIPP), interviewers were asked to keep a systematic record of their mileage and time spent in discovering, locating, and following up persons or households that moved. A detailed analysis of this Mover's Cost Study is presented by White and Huang (1982), who (among other things) reported a mover household follow-up rate of 76 percent (with an eligible person interview rate of 92 percent in interviewed households) and a cost increase of approximately 8 percent attributable to following movers. Of particular interest to potential policy analysts of these data is that nearly 78 percent of the 1979 Research

Panel Wave 6 sample households had never moved. Relative to other longitudinal databases where movers are not followed, then, sample attrition due to this factor is clearly lower in the 1979 Research Panel, and estimates involving variables related to residential mobility less subject to such nonresponse bias.

CONCLUSION

In conclusion, from this brief overview it should be clear that, aside from their great analytic potential, some of these tests and experiments may also have a deletorious or confounding impact on certain substantive analyses that might be conducted using data from the 1979 Research Panel. And, in a few cases, these methodological have some positive implications for such analyses as well. Such implications range from some obvious deficiencies in some of these data highlighted by these field tests to more subtle influences on data quality and variances due to the experimental treatments imposed on the survey design. Especially with regard to the latter, the positive benefit of including such methodological tests in the survey design is that the potential influence of such factors on substantive results from this survey may be directly assessed in data analyses. It is important to note, however, that if these factors are not so examined their influence may lead to distorted or spurious conclusions. By describing some possible road blocks that these tests and experiments may throw in the way for substantive analysis and interpretation, this paper has sought to illustrate the need for both consumers and analysts of these data to keep their methodological nature clearly in mind and, where possible, to assess directly the potential influence of these factors on research results.

REFERENCES

Czajka, John L.
1982 "How complete was food stamp reporting in Wave II of the ISDP Panel." Manuscript. Washington, D.C.: Mathematica Policy Research, Inc.

Czajka, John L.
1983 "Subannual income estimation." Pp. 87-97 in Martin H. David (Ed.), Technical, Conceptul, and Administrative Lessons of the Income Survey Development Program (ISDP). New York: Social Science Research Council.

Coder, John F.
1980 "Some results from the 1979 Income Survey Development Program Research Panel." Pp. 540-545 in Proceedings of the Section on Survey Research Methods, American Statistical Association.

David, Martin H.
1983 "Measuring income and program participation." Pp. 3-20 in Martin H. David (Ed.), Technical, Conceptual, and Administrative Lessons of the Income Survey Development Program (ISDP). New York: Social Science Research Council.

Kaluzny, Richard
1982 Evaluation of Experimental Effects, 1979 Research Panel--Wave 1. Final Report. Princeton, NJ.: Mathematica Policy Research, Inc.

Kaluzny, Richard
1982 Evaluation of Experimental Effects, 1979 Research Panel--Wave 2. Draft Final Report. Princeton, NJ.: Mathematica Policy Research, Inc.

Kasprzyk, Daniel
1983a "Some research issues for the Survey of Income and Program Participation." Pp. 686-691 in Proceedings of the Section on Survey Research Methods, American Statistical Association.

Kaspryzk, Daniel
1983b "Social security number reporting, the use of administrative records, and the multiple frame design in the Income Survey Development Program." Pp. 123-141 in Martin H. David (Ed.), Technical, Conceptual, and Administrative Lessons of the Income Survey Development Program (ISDP). New York: Social Science Research Council.

Kulka, Richard A.
1983 "Tests and Experiments." Pp. 4-1 to 4-39 in P. Nileen Hunt, et al., ISDP 1979 Research Panel Documentation. Springfield, VA: National Technical Information Services.

Moore, Jeffrey C., and Daniel Kasprzyk
1984 "Month-to-month income recipiency changes in the ISDP." Paper presented at the annual meetings of the American Statistical Association, Section on Survey Research Methods, Philadelphia, August 13-16.

Roman, Anthony M., and Diane V. O'Brien
1984 "Findings from the student follow-up investigation of the 1979 Income Survey Development Program." Paper presented at the annual meetings of the American Statistical Association, Section on Survey Research Methods, Philadelphia, August 13-16.

Vaughan, D.R., and C.G. Lancaster
1980 "Applying a cardinal measurment model to normative assessments of income: Synopsis of a preliminary look." Pp. 546-551 in Proceedings of the Section on Survey Research Methods, American Statistical Association.

White, Glenn D., Jr., and Hertz Huang
1982 "Mover follow-up costs for the Income Survey Development Program." Pp. 376-381 in Proceedings of the Section on Survey Research Methods, American Statistical Association.

Whiteman, T. Cameron
1983 "Lessons to be learned from ISDP: The measurement of nonfarm self-employment income." Pp. 73-82 in Martin H. David (Ed.), Technical, Conceptual, and Administrative Lessons of the Income Survey Development Program (ISDP). New York: Social Science Research Council.

Ycas, Martynas, and Charles A. Lininger
1981 "The Income Survey Development Program: Design features and initial findings." Social Security Bulletin 44 (November): 13-19.

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

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

Introduction: Need for a Longitudinal Survey.

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

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

Implementation of a Longitudinal Survey.

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

The shift from an address sample for the first visit to a person sample in subsequent visits presented unique challenges to the planning staff, regional office staffs, and interviewers. Updating procedures for the address listings, noninterview classifications, interviewing procedures, and many other activities required for surveys maintaining an address sample were not appropriate. New controls and follow-up procedures, some requiring interregional office cooperation, were implemented. Interviewers received extensive training on new noninterview classifications and movers' procedures. Office staff maintained extensive clerical controls to guarantee the receipt of control cards and questionnaires from interviewers and to monitor the processing of over 40,000 person records that were uniquely identified.

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

Wave 1 Address Sample Procedures.

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

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

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

Development of Special Procedures for the Longitudinal Survey.

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

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

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

(1) Persons who join Wave 1 sample persons in later waves are not followed to new addresses unless these additional persons remain with Wave 1 sample persons who are 15 years old or over;

(2) Persons who move out of the sample universe are not followed. These are persons who become institutionalized, move outside of the United States or live in an Armed Forces barracks;

(3) Children under 15 who move and are not accompanied by a sample person who is 15 years old or over are not followed.

The geographic area covered by personal visit follow-up is extensive. Based upon the 1980 census population distribution, about 130 million persons live in areas within SIPP PSUs; another 87 million persons live within 100 miles of the outer boundary of a SIPP PSU. We counted 226 million persons in 1980; 217 million are within our currently covered areas--96 percent of the population.

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

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

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

The SIPP Identification System.

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

The components of the operational SIPP identification system are:

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

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

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

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

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

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

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

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

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

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

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

(1) The portion of the control number consisting of PSU, segment, and serial number is similar to the numbering system used by other major surveys conducted by the Bureau.

(2) Interviewers are able to assign person numbers during the course of the interview. The person number is used in various parts of the questionnaire during the interview. This number is also transcribed to several other survey documents during the interview and immediately afterward during clerical coding operations. A person number assigned after the time of interview does not provide this immediate linkage.

(3) The person number itself has relatively few digits, reducing the possibility of transcription errors.

Several disadvantages have been noted:

(1) Duplications of person numbers for additional persons (persons added after Wave 1) can conceivably occur in situations where households have split and are in different regional office jurisdictions. The computer processing system identifies these duplicates and the regional office staff corrects them during processing.

(2) Mergers between two separate sample households require special procedures. If this situation occurs, one set of controls is retained for the merged household. New person numbers are assigned to those persons who lose their original identifiers. Interviewers record both the old and new ID numbers on the control card to provide a means for linking the two ID's. By the end of the second Wave, this had occurred once.

Monthly Cross-Sectional Households.

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

Entry 1 - birth)
2 - marriage)
3 - other)
4 - 5/)

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

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

Clerical Field Controls.

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

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

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

(1) During the keying operation all persons listed on the control card who are 15 or over and are current household members must have an accompanying questionnaire. This check is done automatically at each keying station. Keying is done in the regional offices and immediate resolution of missing questionnaires is required.

(2) At the end of each of the four months of each wave, a centralized check-in is completed in Washington. A control card record must be transmitted for every person showing an active status on a master file maintained in Washington of all active records. Offices cannot close out an interview month until every active status person is accounted for and some demographic data--age, race and sex--is verified to make sure that we are not checking in the wrong person. Each missing case is referred to the appropriate regional office for resolution.

Experience with Following Movers.

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

(1) Percentage of movers found: about 80%.

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

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

Recommendations for future SIPP Panels.

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

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

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

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

As described earlier, in the 1984 SIPP, only persons who are 15 or over are followed to new addresses; sample persons who are under 15 years old are not followed unless they move with a sample person who is 15 or over. However once they become 15 they are eligible for interview along with other members of their households. They are missed in the 1984 panel if they move before turning 15 and are not accompanied by a sample person who is 15 years old or older. Their absence may result in some bias in the survey data. In future SIPP panels, all sample persons who are 12 years old or older at the time of the first interview will be eligible for follow-up. When a person who was 12 years old at the time of the first interview moves by him- or herself to a new address, occupants of the new household will be interviewed according to standard procedures--that is persons 15 years old and over will be administered a questionnaire. When the sample person turns 15, that person will also be administered a questionnaire.

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

(1) Reassigning Wave 1 eligible noninterviews in Wave 2. Interviewers will be provided with instructions for obtaining household rosters and assigning person numbers retrospectively--i.e., as of a date approximately four months prior to the date of the second interview.

(2) Adjusting the computerized check-in system to allow for new serial numbers (representing persons or addresses) to be introduced in Wave 2. This will provide flexibility for including missed Wave 1 housing units.

(3) Developing a questionnaire that is appropriate for telephone interviews. This could be administered to persons who are not followed for a personal visit.

(4) Increased automation over the next few years will eliminate much of the time consuming clerical operations associated with the check-in, control and monitoring of assignments.

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

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

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

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

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

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

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

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

MANAGING THE DATA FROM THE 1979 INCOME SURVEY DEVELOPMENT PROGRAM

Pat Doyle, Mathematica Policy Research
Constance F. Citro, National Academy of Sciences

During 1979 and 1980 the Department of Health and Human Services and the Bureau of the Census, with support from other federal government agencies including the Food and Nutrition Service, USDA, administered a panel study of households representative of the civilian noninstitutionalized population in the United States called the 1979 Income Survey Development Program (ISDP) Research Panel. The survey was designed as the final pretest for the Survey of Income and Program Participation (SIPP) which had been under development since 1975 and was fully implemented in late 1983. The 1979 panel study was extremely complex due to the efforts put forth to improve the measurement of income, net worth, and program participation and to increase the information available on behavior, attitudes, expenses and disposable income of the population.

The complexity of the 1979 ISDP survey design led to the production of public use files which are cumbersome to use thus making it difficult to access the newly available data for research. The subject of this paper is to describe a project conducted by Mathematica Policy Research (MPR) under contract to the Food and Nutrition Service, USDA, to solve the data access problems through the use of data base management system technology. The DBMS chosen for this work was RAMIS II™ developed and distributed by Mathematica Products Group. The system developed by MPR is referred to as the ISDP/RAMIS II system.

It is important to note that a number of problems that were confronted in designing the access system described in this paper have been resolved in the release of the public use ISDP files (in fact, data from the ISDP/RAMIS II system were the source of some of these improvements). Furthermore, some, but by no means all, of these access problems have been explicitly taken care of in the design of the SIPP. Consequently, designing an access system for the new survey should be easier than for the ISDP. It is also true that the best design for a SIPP access system is likely not to be the design chosen for the ISDP system.

In the subsequent section, an overview of the panel study with emphasis on the contents and problems of the data files is provided. The report concludes with an overview of the newly created system with a summary of the data problems solved in the course of this work. For detailed information on the contents and use of the ISDP system, the reader is referred to Doyle and Citro (1984).

Overview of the ISDP and Its Applications

Figure 1 gives a graphic representation of the key features of the ISDP design. Briefly, note that:

- -- There were 6 waves of interviewing providing 12 to 15 months of data for each household.
- -- Interviewing was staggered; one-third of the sample was interviewed each month, with, thus, a different 3-month reference period for each rotation group.
- -- This pattern was regular, except that the third rotation group, for various reasons, was skipped over Wave 4.
- -- Each wave asked a core set of items, including monthly income and employment, plus one-time supplemental items.

The SIPP design for the first panel is very similar, including skipping one wave for part of the sample.

The ISDP, by virtue of gathering detailed month-by-month data over a span of at least a year, offered the potential for exciting research that simply could not be carried out before. But, to make it possible for the researchers at MPR to realize that potential, we had to design an access system that would do the following:

- -- Generate reports and analysis files from individual waves, undoubtedly the easiest way of using the data
- -- Generate reports and analysis files linking data across waves
- -- Let researchers apply different rules to identify households and families across waves for longitudinal analysis
- -- Link supplemental data collected in one wave to core data in other waves
- -- Make it possible to carry out sophisticated statistical as well as tabular analysis of the data
- -- Make it possible to use the ISDP data with data from other sources, for example, 1980 census summary data.

All of these access requirements apply equally well to the SIPP.

Problems for Access Posed by the ISDP

Various design features of the ISDP posed more or less serious problems for developing an access system that would satisfy the requirements just listed. These are summarized below.

- o Staggered Interviewing. The use of a staggered interviewing schedule results in a situation where data from more than one interview must be accessed to study a common calendar period for the entire sample (except where the user can make do with the single calendar month that is common to all rotation groups within a wave).
- o Skipping Wave 4. The alteration of the interviewing schedule to have the third rotation group skip over the Wave 4 interview means that, although two-thirds of the

sample cases have a full 15 months of data (from the five regular waves if they did not attrite), the other third has only 12 months. Moreover, the third rotation group does not have responses to any of the topical supplemental items asked at Wave 4.

o Different reference periods for wave-specific information. For any one interview, there is a potential mismatch between the wave-specific data and the monthly data, given that monthly data for the month of an interview were actually asked at the subsequent interview.

o Identifier problems. The Census Bureau encountered problems in uniquely identifying individuals across the survey waves, necessitating creation of a new unique person identifier, called the link index, as a separate file from the interview data files. It also turned out that the Bureau erroneously included some persons on the cross-section interview files who were not in fact present and vice versa.

o One-time wave-specific supplemental data. The fact that important data were asked on a supplemental one-time basis creates problems for using these items together with the monthly and quarterly data.

o School lunch data problems. The ISDP files include valid data only for the last child in a family, and these data were erroneously written into the records for all other children.

o Lack of editing on Wave 6. In the case of Waves 1-5, the Census Bureau performed edits on demographic variables and also edited income recipiency flags. No editing was performed on the Wave 6 data, which were collected in an entirely different format.

o Asset income reporting experiment. This experiment creates practical problems of associating asset income data with other data for each month of the panel.

o Incomplete determination of monthly unit composition. The design of the cross-section files, coupled with a high level of noise in the data on arrival and departure dates, made it very difficult to assemble a stream of monthly unit composition indicators consistent with reported monthly economic data.

o Absence of longitudinal weights and imputations for missing data. The cross-section interview files contain weights and also imputations for missing income and employment data that were constructed strictly on a cross-section basis which are not suitable longitudinal studies.

o Absence of longitudinal editing. With the exception of editing age and sex in the construction of the unique identifiers, no longitudinal edits were performed on the demographic variables.

These characteristics of the ISDP survey make retrieval of the information for analysis cumbersome and expensive. This is particularly true for longitudinal applications of the data such as the study of turnover in the Food Stamp Program.

The difficulty in using the ISDP for research was compounded by the structure of the available data files. At the time this project was carried out, the most suitable input file was a concatenation of cross-section files from all five waves. The format for each cross-section was similar to the public working files currently available (NTIS, 1982) except that the family level had not been fully developed. The records from all five waves were grouped by PSUSERIAL and a level 1 record was created which recorded information common to each group such as rotation.[1] In addition to inserting the level 1 record, the Bureau also merged the link index (constructed unique person identifier) and longitudinal edited values of age and sex to this file. However, the Bureau deleted from this file the results of the cross-sectional imputations for income and employment data. The rationale for this omission was the unsuitability of these imputations for longitudinal analysis, the purpose of the concatenated file.

This file was extremely cumbersome to access due to the lack of a true hierarchical structure, the large number of different record types (data from each topical module were recorded on a separate record with a distinct record length and layout) and the fact that some of the newly created person identifiers were erroneous.

Overview of the ISDP/RAMIS II System

The objective of this data base development effort, as noted above, was to take the information available on the series of cross-section files described above and array it in a manner that would facilitate longitudinal as well as cross-sectional analysis. The results of this effort were two RAMIS II data bases, one called SIPPMASTER and one called MH for monthly households. SIPPMASTER is the main file in that all of the data collected during each wave are stored there. This file is used for all cross-section applications as well as longitudinal applications which do not involve the formation of longitudinal households or other groupings of individuals. The MH file is the data base designed to support the construction of longitudinal units. It essentially provides information on monthly household, family, and food stamp unit composition. The data in MH are arrayed to permit a user to develop a definition of longitudinality and apply that in the construction of a longitudinal unit file. Once the longitudinal unit itself is determined, the user can employ the data stored in SIPPMASTER to derive variables like total household monthly income which reflect the longitudinal unit characteristics.

The remainder of this section provides an overview of the contents of the ISDP/RAMIS II system. A detailed discussion of the motivation for choosing this file design and the procedures required to develop this data base is described in Doyle and Citro (1982).

SIPPMASTER. Figure 2 displays the logical organization of SIPPMASTER. It has a hierarchical structure with fifteen levels, five of which are real and ten of which are virtual.[2] The five real levels are wave, household, family, person and month. Some relevant comments on each of these levels follow:

Wave. (Level 1) Indicators for Waves 1 through 5 are contained in SIPPMASTER on level 1. The data from Wave 6 are treated as supplemental and are therefore stored in the virtual level PM (level 7). SIPPMASTER is physically separated into 5 data bases, one for each wave. They are linked together with RAMIS II USE commands to logically form one data base.

Household. (Level 2) This reflects household composition at the time of the interview. The household identifier (HHID) uniquely identifies households within wave. It cannot be used to identify households longitudinally. Non-interview households in each wave have entries at this level, however data for all other levels are zero. The contents of the household level consist of the data found in the household record in the cross-section files prepared by the Census Bureau.

Family. (Level 4) The family level simply identifies family units within households as they existed at the time of each interview. Primary individuals, secondary individuals, and outmovers are treated as one person families.

Person. (Level 5) This contains interview specific data for each individual and retrospective data that were not collected for specific calendar months such as total weeks unemployed. The identifier for level 5 is the link index (called PERID in RAMIS II) so that each person sampled is identified in the same way across all waves. The data for the person level were derived from record type 5 of the cross-section files. Some relevant points: outmovers in a given wave are included for that wave but have 0 in the weight fields; the weights are cross-sectional; all person identifiers with values exceeding 200000 should be deleted for longitudinal analysis but not for cross sectional analysis; corrected age (CORAGE) should be used instead of edited age (AGEED) except that corrected age is 0 on Wave 2: income recipiency flags on level 5 are not to be used to determine item nonresponse as they were retained here for other reasons (for example, if the interest flag in Wave 1 is 1 on level 5 but there is no entry for that income type in the WU or MU associated files, then the person was reported to have had an interest producing asset but did not actually receive interest income during the Wave 1 reference period).

Month. (Level 12) This represents the reference period for each wave. All months in the survey have been numbered longitudinally so that, for example, the 3 months pertaining to Wave 2 are 4, 5, and 6. Aside from identifying the longitudinal reference months, this level contains numerous fillers intended to support the construction of longitudinal household (or other aggregate unit) files.

The remaining data available through SIPPMASTER are stored in associated files which can be accessed directly if desired. A summary of the contents of each can be found in Doyle and Citro (1984).

MH. Figure 3 describes the logical organization of MH. It is a relatively simple hierarchical file with five real levels and one virtual level. This file reflects the outcome of a complicated procedure designed to determine monthly household and food stamp unit composition from the data collected in the 1979 ISDP Research Panel. Documentation on the methodology employed in the development of this file is included in (Doyle and Citro, 1984). The contents of this file are described below followed by a section summarizing how it is used to develop longitudinal units.

Unlike SIPPMASTER, MH contains a limited number of variables. It is comprised mostly of pointers detailing who lived with whom during each month covered by the first five waves of the survey. The remaining variables provide descriptive characteristics such as age and relationship to reference person which are necessary to effectively determine longitudinal units. Each of the levels of MH is described below.

PSUSERIAL. (Level 1) This level contains the scrambled values PSUSERIAL as well as the rotation group identifier. For the ISDP all persons who ever resided together have common values of PSUSERIAL, so this level was created to increase the efficiency of data retrieval and to minimize storage costs.

MONTH. (Level 2) This level simply identifies the month. Longitudinal reference months as described for SIPPMASTER were used. For rotation groups 1 and 2, the months range from 1 to 16 and for rotation group 3 they range from 1 to 13. Note that household composition can be described for one more month than is covered by the retrospective data collected in the ISDP. This extra time period is the month of the final interview.

Household. (Level 3) This level describes who lived with whom during each month and the Food Stamp Program participation and benefits of that group. The contents are the monthly household identifier and food stamp recipiency and amount variables for up to two food stamp units.

Family. (Level 4) This is an indication of family groupings within monthly households. The contents are family identifier, family type, and family kind.

Person. (Level 5) This level contains an entry for each person for every month he or she was present in the sample. The key to this level is PERID, the same identifier used in SIPPMASTER. The other variables stored in this level are age, relationship to reference person, marital status, food stamp unit membership and variables necessary to link to SIPPMASTER.

PD. (Level 6) This is a virtual level in MH. The associated file is called PD and it is the same PD file accessed through level 6 of SIPPMASTER. It contains presence in sample indicators as well as constant demographic data such as sex.

The intended use of the MH data base is to determine longitudinal units. In developing the ISDP/RAMIS II system, one objective was to allow researchers flexibility in the development of the definition of what constitutes the same unit when viewed over time. For some applications it may be appropriate to define a unit as being the same from one month to the next if all adults remain the same. For another application it may be sufficient to only require that the reference person (household head) be the same. More complicated definitions may be required in other situations. An example might be that units are the same if the composition changes from one month to the next are restricted to birth of a child, loss of a peripheral adult, e.g., an older daughter leaves for college, or a death of one spouse in a husband-wife primary family.

Each of these three definitions can be specified with the ISDP/RAMIS II system as can many others. The procedure is as follows. Using the preferred definition, an algorithm for uniquely identifying each unit each month is developed. In the second example above, this would simply involve assigning the PERID of the reference person to the monthly unit as the identifier. Next, a comparison across months within PSUSERIAL groups is made. All monthly units with common values of the newly created identifier constitute one longitudinal unit. Finally, an extract is created which records the available information organized by the longitudinal unit identifier.

The available data from MH are primarily demographic, the exception being Food Stamp Program characteristics. The user will of course also desire economic data to support the analysis of the longitudinal units. This can be achieved through the extraction of data from SIPPMASTER.

Conclusion

This paper describes a system to access data from a complicated longitudinal survey of households when the survey itself was in its development stages. It represents a successful attempt to apply modern DBMS technology to solve access problems posed by complex social science data collection efforts. Some of its features are:

- Good report generation for easy specification of tables and extracts
- Procedural language interface to allow the use of FORTRAN or PL/1 to conduct complex applications
- SAS interface to permit more sophisticated statistical analysis.

The system is, of course, not without drawbacks. For example, the hierarchy imposed in the primary file, SIPPMASTER, is cumbersome and, with recent developments in relational data base technology, unnecessary. This structure could easily be simplified today. Furthermore, the system is on-line and therefore require large amounts of disk storage. As the cost of mass storage goes down with improved hardware now being developed, this will become less of a problem.

In spite of these imperfections the ISDP/RAMIS II system works. It represents the first truly integrated ISDP data base available to the public for research. With this system users can and indeed have carried out analyses that truly exploit the longitudinal nature of the data.

FOOTNOTES

[1]On the publicly available data bases, PSUSERIAL is a nine character field which uniquely identifies all households in Wave 1. Together with person number it was originally intended to uniquely identify persons followed in the panel.

[2]A virtual level is a level for which the data are not physically stored in the file. Instead there is an internal record of the location of another file which contains the information. With a DBMS, this other (or associated) file is accessed automatically when data from it are requested.

REFERENCES

Carr, Timothy; Doyle, Pat; and Lubitz, Irene. "An Analysis of Turnover in the Food Stamp Program." Draft Report. Washington, D.C.: Mathematica Policy Research, 1983.

Doyle, Pat and Citro, Constance F. "The ISDP/RAMIS II System and Its Development." Draft Report. Washington, D.C.: Mathematica Policy Research, 1984.

Doyle, Pat and Citro, Constance F. "Proposed Design Strategy for Storing and Accessing Data from the 1979 Income Survey Development Program Research Panel." Draft Report. Washington, D.C.: Mathematica Policy Research, 1982.

National Technical Information Service. Income Survey Development Program: 1979 Research Panel Documentation. Springfield, V.A.: U.S. Department of Commerce, 1982.

FIGURE 1

Survey Waves -- 1979 ISDP Research Panel

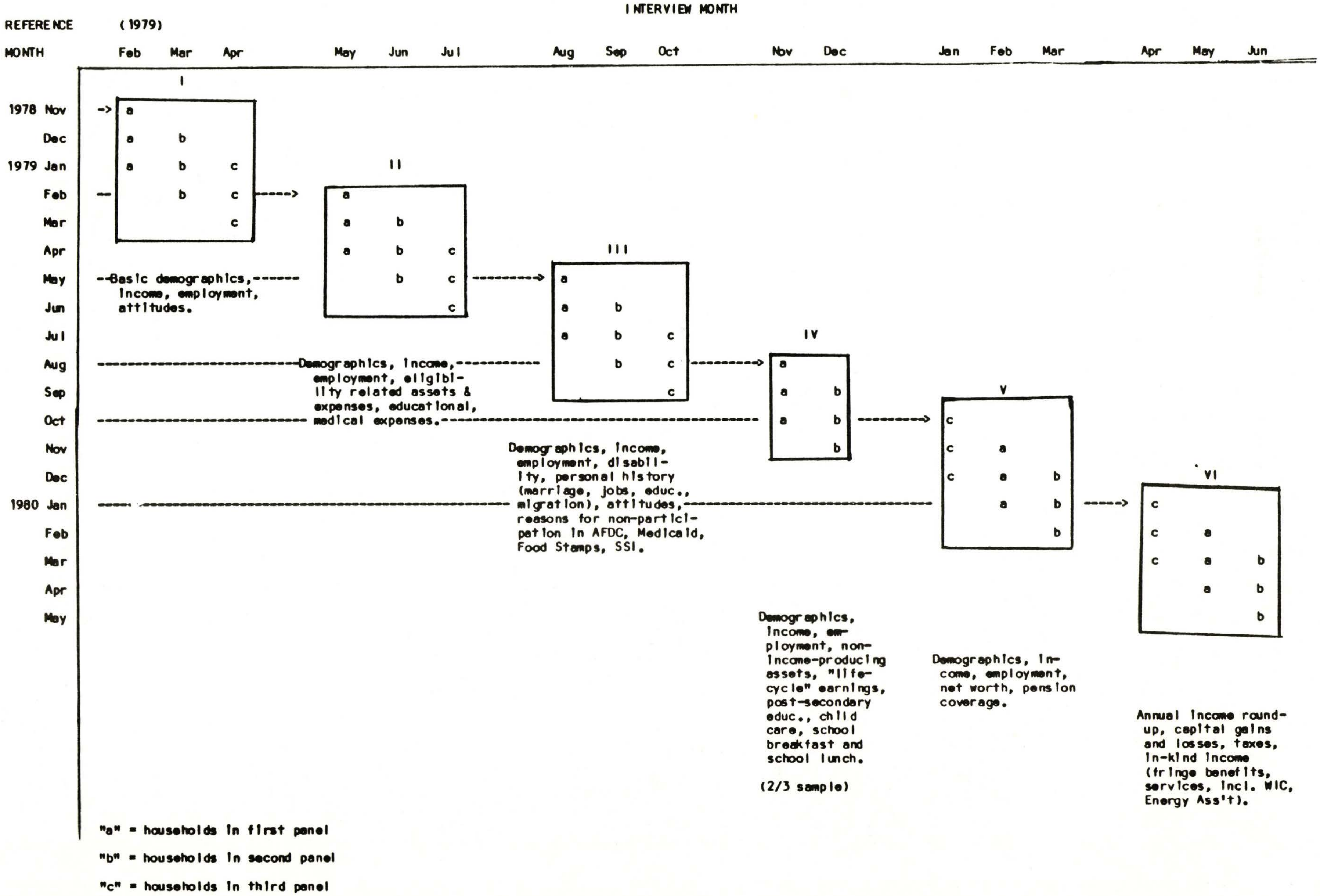


FIGURE 2

FILE STRUCTURE FOR SIPPMASTER

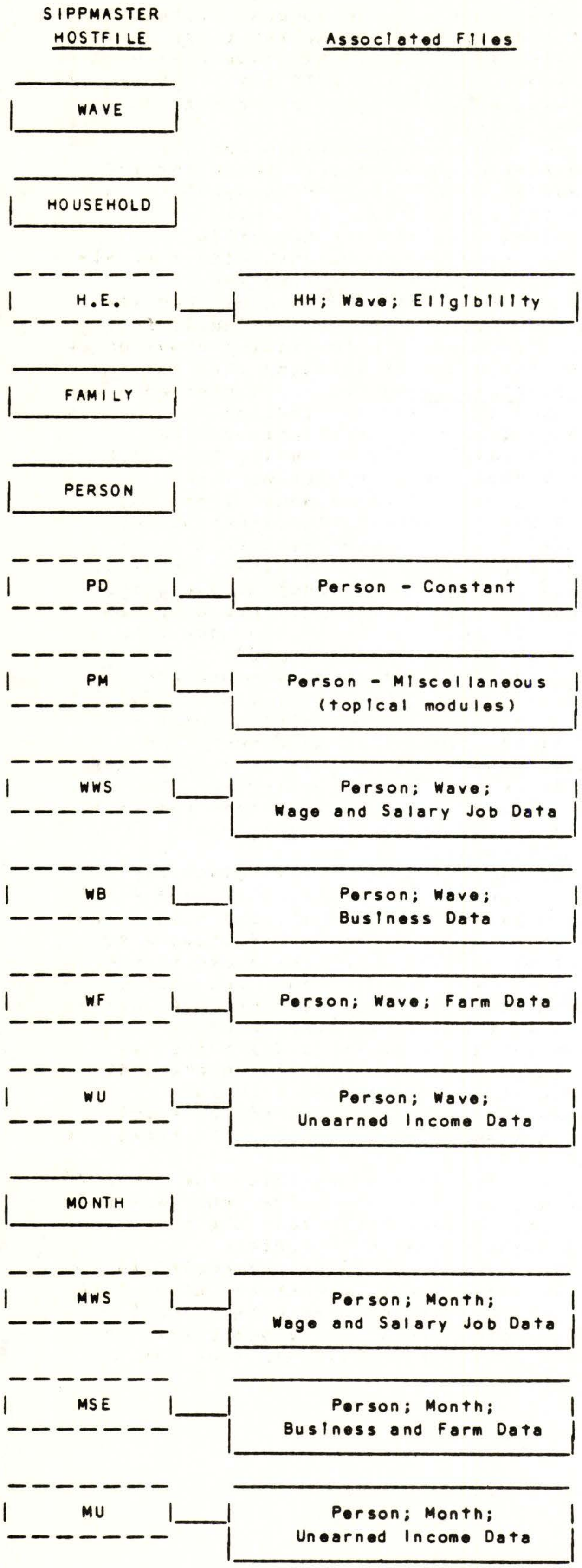
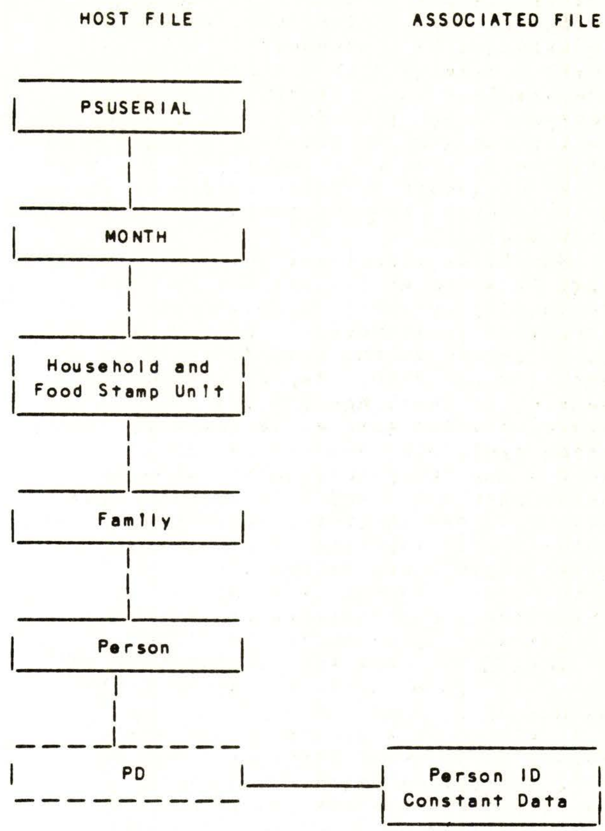


FIGURE 3

RAMIS II FILE REFLECTING MONTHLY UNIT COMPOSITION (MH)



DISCUSSION

Greg J. Duncan, Survey Research Center

The papers of Jean and McArthur and Doyle and Citro focus on several aspects of data collection and management associated with the Survey of Income and Program Participation (SIPP) and its elaborate pretest, the 1979 Income Survey Development Program (ISDP). Both must address issues arising from the basic design of longitudinal surveys of individuals and households and it is worth beginning with a brief review of the sampling theory behind the SIPP design. Since this design is so similar to the one used in the longitudinal study with which I am most familiar--the Panel Study of Income Dynamics (PSID)--I will draw heavily from the 17-year history of that study.

Many cross-sectional surveys obtain their samples of individuals and households by sampling dwellings. Longitudinal surveys can do this as well, as evidenced by the procedures of both the SIPP and PSID. Representative samples of dwellings provide representative samples of subunits within those dwellings--households, families, Food Stamp recipiency units, AFDC recipiency units and individuals. The selection probabilities of each of these subunits are identical to the selection probabilities of the dwelling itself. With a properly specified set of rules regarding the definition of units and the tracking of those units over time, a longitudinal study such as the SIPP or the PSID can maintain a representative sample of each of the various subunits over time. This requires that newly formed subunits of interest (families, AFDC recipiency units, etc.) enter into the sample with known selection probabilities in order to reflect corresponding changes that are taking place in the population at large. It requires that individuals be classified as either "sample" or "nonsample" and that explicit rules be followed consistently in the event of dramatic changes in the composition of units. In the SIPP, as in the PSID, for example, nonsample individuals who join the sample through marriage are followed only as long as they remain attached to a household containing a sample member. Once they regain their independence from all sample members, they are no longer followed.

In general, these sampling considerations require that the study have good systems for (1) tracking all sample individuals, regardless of where they go, (2) allowing individuals to join the sample to provide accurate information about the household in which sample individuals reside, (3) having a fool-proof system of identification numbers for all individuals, sample and nonsample, and (4) storing the data for the individuals and aggregations of individuals so that an analyst can perform a variety of analyses on these individuals in an efficient way.

Most of the field control procedures outlined by Jean and McArthur are quite similar to the ones that have been used successfully by the PSID for 17 years. I do have a few comments about some of them, however.

1. Not all individuals who are institutionalized appear to be carried along as a part of SIPP households. In the PSID, individuals who are institutionalized and cannot be interviewed are associated with a sample family for as long as they remain institutionalized. Of course they are not considered part of the family for most purposes, but the family provides us with the means of tracking them and then reestablishing contact with them when and if they leave the institution. It may be tempting to drop institutionalized individuals from the sample, but there are a substantial number of them, especially at younger adult ages. A strategy of dropping institutionalized individuals in a country with a compulsory, universal military service, would result in all young people being dropped from the sample! Not keeping track of young children who move into institutionalized housing of various types or with relatives who are not sample members means that the SIPP will be unable to inform analysts about such children. (The PSID does not follow these young children either.) They may be too expensive to follow, but the decision of not following them should be based on an appreciation of the consequences.

2. Model-based statisticians may not appreciate the distinction between sample and nonsample individuals and will lament the fact that nonsample individuals are dropped by SIPP once they leave sample households. The PSID does not follow nonsample individuals either, but perhaps this is a mistake. Some methodological work conducted by Finis Welch and his colleagues on the PSID has detected no significant differences between behavioral models estimated for sample and nonsample individuals. (Becketti, et al. 1983.)

3. The nonresponse rules for the SIPP are not entirely clear from the Jean and McArthur paper, especially the rules regarding attempts to contact nonrespondents to waves subsequent to the first one. The PSID does not attempt to recontact these nonrespondents and I think that that is the biggest flaw in the PSID design. Evidence from the new

youth cohorts of the NLS indicates that nonrespondents in one wave are often quite willing to respond to subsequent waves. One gets the impression that refusals or contact difficulties are often quite transitory in nature.

4. The Jean and McArthur paper mentions but does not emphasize the importance of obtaining the name, address and phone number of a contact person who might know the whereabouts of sample households if they move. More conventional means of following individuals such as through forwarding addresses sometimes do not work precisely because the individuals do not wish to be followed easily. In our experience, the contact information is invaluable.

5. Telephone interviewing is mentioned as a possible way of preserving high response rates. The PSID experience suggests that this is indeed true and that data quality does not suffer unduly from switching interviewing modes. Indeed, a substantial number of recontact calls are made to PSID respondents to clean up unclear interview information. Telephones also provide a means of not only reaching geographically remote respondents but also respondents whose time schedules make telephone interviews much more likely to succeed.

Before turning to the subject of the Doyle and Citro paper I would like to make a comment on the interaction between the data collection and data management. Too often we compartmentalize the two without realizing how intimately they are related. As illustrated in the Doyle and Citro paper, data analysts often discover apparent inconsistencies or outright errors and are in the worst position to make an informed judgement about the problems. Data collectors ought to anticipate problems of this sort and have significant resources allocated to solving them. Most of the problems must be resolved by returning to the original protocols, at least briefly, to understand the nature of the problem.

What now of the data structure and methods proposed in the Doyle and Citro paper? Several basic questions come to mind.

1. The most basic question to be asked of any proposed data structure is "Is it feasible?" That the proposed structure has been used with success for several ISDP projects suggests an affirmative answer to this question.

2. The second question, more difficult to answer from the information contained in the paper, is "Is it efficient?" or, more properly stated, "Under what circumstances is it efficient?" Does one need a dedicated machine capable of grinding away throughout the night to select an abstract from the data set with this system, or is it feasible to use the proposed system in a computer environment in which CPU is priced at its marginal cost? I suspect that the proposed system is not very efficient in the latter type of computing environment but I could not tell from the information contained in the paper.

3. Since most "computing" costs are the labor costs of the programmers and other analysts rather than the machine charges, the third question is "Is it easy to use?" Apparently once one has acquired a great deal of specific training about the proposed system, it is fairly straightforward. But outside analysts are encouraged to consider avoiding the data abstracting complications by delegating that work to those who are more familiar with the system.

The data structure that is proposed is modelled after the exceedingly complex file structure used by the Census Bureau. Surely there is a simpler method than an eight-level hierarchy for each wave and four files each with a fifteen-level hierarchy and a completely separate six-level hierarchy that can be used to sort out different aggregations of individuals. The PSID files are more complicated in that they have more waves of data but are simpler in that they are in only one aggregation--the family. It has but two levels to its hierarchy--the family history and the individual. The term "family history" is chosen with care because a major insight, obvious now but not during the first twelve years of the study, are the data structure implications that stem from the fact that not all individuals in a given family in the most recent wave share the same "family history". In fact, we have about seven thousand current families but over nine thousand family histories. The first level of the hierarchy, then, is the family history; the second level consists of the individual histories of all of the individuals who share that same family history. One could also construct "household histories", "Food Stamp recipiency histories", etc. as additional hierarchical levels or as separate records in a networking data structure. These simpler hierarchies require that some of the information from the individual data record be aggregated into the family or household record and this work is probably best done at the Census Bureau rather than having outside analysts attempting to do this with the information they have at their disposal.

A final comment concerns a limitation again attributable to the way in which the Census Bureau processes its data rather than to the organizations such as Mathematica Policy Research that attempt to make sense out of it. Implicit in the file structure is the assumed need to aggregate individuals into households or other sensible units, but not the possible need to relate individuals to one another. One could think of a file in which all sample individuals had data records that contained information on all individuals who had been or were about to be related to them in some way (by blood, marriage, adoption, or sharing the same dwelling). Information on the related individuals would include wave by wave (or, in the case of SIPP and ISDP, month by month) information on how the individuals were related and whether they shared the same dwelling, family, household, Food Stamp recipiency unit, etc. For most purposes this would be the most general file structure for SIPP, enabling the analysts to distinguish step children from natural children, ex-spouses, and other relatives so that one could analyze the economic consequences of divorce, etc. This would, of course, require a great deal more information than is now currently provided in the Census Bureau's current "relation to head" coding. But the added detail would enable the construction of a file structure that would be of greatest use.

References

Becketti, Sean; Gould, William; Lillard, Lee and Welch, Finis Attrition from the PSID, Santa Monica, CA: Unicon Research Corporation, November, 1983.

DISCUSSION

Richard C. Rockwell

The papers by J. C. Moore and D. Kasprzyk, A. M. Roman and D. V. O'Brien, and R. A. Kulka are important efforts to examine sources of error other than that due to sampling. I would hope that this concern with nonsampling sources of variance continues throughout the program of evaluation and research on the new SIPP, and that the results are reflected in reports coming out of the SIPP. There is ample reason to suspect that, in a survey as complex and difficult as is the SIPP, error due to sampling will be swamped by error due to miscommunication between respondent and interviewer, respondent dissembling, respondent ignorance, etc.

The Moore and Kasprzyk paper tackles the problem that the ISDP (the pilot study for the SIPP) can be seen as having measured more change between waves than within waves. The authors argue persuasively that this cannot be reflective of true conditions, citing a number of reasons for thinking that not only did respondents make errors that resulted in the appearance of change between waves but also that post-interview processing may have contributed significantly to this phenomenon.

Their analysis heavily depends on the correct matching of persons across waves. An earlier analysis, by Graham Kalton and James Lepkowski, had depended on a file that contained many identifiers with erroneous codes, making matches very prone to error. Moore and Kasprzyk used instead a "definitive" data file produced by Mathematica, which linked data for the first five of the six waves. The authors assert that this "apparently" corrected the matching problems by correcting person identifiers. This is a curiously ambiguous way of describing what must be one of the central assumptions of the analysis. If it is not now possible for the Bureau of the Census to construct its own matched file, then I suggest that a second-best alternative is to document the matching process used by Mathematica and to obtain an independent validation of that process. This is particularly important given the research now beginning on the ISDP, with support from NSF and other agencies.

The authors do point out the intriguing possibility that the matched file processed by Kalton and Lepkowski contained a higher proportion of correct matches, although fewer matches in all than does the Mathematica file. Reporting that even a low rate of mismatching can produce the level of between-wave change observed (on the basis of a computer simulation), Moore and Kasprzyk raise the question that post-interview processing may have played a large role in the results obtained.

The paper points indirectly to one of the big methodological questions facing the SIPP: when data are not only collected longitudinally (as in the CPS and the SIPP), but are also to be analyzed longitudinally (as in the SIPP), it becomes necessary to examine and perhaps rethink established procedures: editing for consistency within records, imputation for item nonresponse, substitution of persons for missing responses, sample weighting across time, etc. Despite years of experience with the PSID and the NLS, for example, there are not yet widely-accepted solutions to such problems. With the SIPP demanding solutions it is imperative to undertake research now.

The Roman and O'Brien paper focuses on one experiment of the ISDP, the comparison of data obtained from college students living away from home with data obtained from proxies, usually their parents. The experiment was conducted during November and December, certainly not the best months to find students resident at school.

Facts, fate, and perhaps a few gremlins took whacks at the sample size. Over one-quarter of the students who were identified were not interviewed, because their school was more than 50 miles distant. Not all parents gave permission for the interview of their students. Not all students were at home. Not all of the completed interviews could be matched with parent interviews. From a potential sample of 443 students identified as usually living away from home, the result is a sample of only 167 matched proxy-student records. One could argue that the failure to match data is the most fundamental form of discrepancy in parent-student comparisons, and in that event the true sample size is somewhat larger, but even with this increase in sample size it is difficult to generalize from the results, with so large a proportion of the sample lost.

The results seem intuitively right: better jobs come to the attention of parents more than low-paying ones; jobs with fewer hours also attract their attention less. So the proxy data are more like the data provided by the students for the big-ticket items. I wonder whether this result is more general: is the income of any lower-earning member of a household, whether a student, an aged relative, or a spouse, less well-reported by the principal earner in the household than is the principal income? Is there a general tendency to underestimate or otherwise misreport the less crucial items in a family budget?

Kulka makes the important point that the 1979 Research Panel of the ISDP was not primarily designed as a substantive data collection instrument but instead as a flexible vehicle for a number of experiments in the technical and operational problems of an income survey. This means not only that the ISDP is a rich data resource for methodological research but also that substantive research must take account of a variety of design effects.

Because funding for ISDP research was terminated in 1982--just when the data sets were becoming available for research--a number of the experiments described by Kulka have been underanalyzed. Kulka's paper raises more questions than can be answered, as a consequence.

The ISDP results that offer some confidence in the data were based on the particular design adopted in the ISDP. As that design was not transplanted to the SIPP, we must not read Kulka's paper as indicative of the quality of data to be derived from the SIPP. The SIPP may be better, or it may not. The same sort of research agenda planned for the ISDP is needed for the SIPP, so that we can have the confidence in the SIPP data that we can now have in the ISDP.

CASE STUDIES IN PANEL

SURVEY DESIGN:

THE INTERNATIONAL EXPERIENCE

SESSION V

This section is comprised of four papers presented in this session which was sponsored by the Section on Social Statistics.

THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

Roger A. Herriot and Daniel Kasprzyk, Bureau of the Census

Introduction

In October 1983, the Bureau of the Census conducted the first interviews of the Survey of Income and Program Participation (SIPP). The SIPP is a nationally representative household survey intended to provide information on all sources of cash and noncash income, eligibility and participation in various government transfer programs, disability, labor force status, assets and liabilities, pension coverage, taxes, and many other items. Data from the survey will provide a multiyear perspective on changes in income, and their relationship to participation in government programs, changes in household composition, and so forth. The purpose of this paper is to review the need for a new survey, briefly describe the research and development work leading up to the SIPP, the survey design features, procedures, and content of the survey. Data products and current survey and research activities will also be described.

The Need for a New Survey

The development of SIPP arose in response to a recognition that the principal source of information on the distribution of household and personal income in the United States, the March Income Supplement of the Current Population Survey (CPS), had limitations which could only be rectified by making substantial changes in the survey instrument and procedures. For example, the CPS:

1) does not measure monthly income flows and month-to-month changes;
2) provides annual income estimates, whereas eligibility for most Federal programs is based on a monthly accounting period;
3) produces estimates of last year's income based on current household membership;
4) does not measure asset holdings and liabilities and does not provide enough measures of categorical information to produce sound estimates of program eligibility; and
5) underestimates income from transfer programs, retirement and disability income, unemployment compensation, and property income.

Because of these limitations, the Income Survey Development Program (ISDP) began. The purpose of the ISDP, authorized in 1975, was to design and prepare for a major new survey, the Survey of Income and Program Participation (SIPP). The development effort was directed by the Office of the Assistant Secretary for Planning and Evaluation in the Department of Health and Human Services and was carried out jointly with the Bureau of the Census, which assisted in the planning and carried out the field work, and the Social Security Administration (SSA), which administers the major cash income security programs.

The ISDP developed methods intended to overcome the three principal shortcomings of the CPS: 1) the underreporting of property income and other irregular sources of income; 2) the underreporting and misclassification of participation in major income security programs and other types of information that people generally find difficult to report accurately (for example, monthly detail on income earned during the year); and 3) the lack of information necessary to analyze program participation and eligibility.

The principal method by which the CPS shortcomings and the new data requirements were to be addressed was the use of a longitudinal survey design. Persons at sample addresses were interviewed about their income and other characteristics for the previous 3 months. They were then recontacted at regular intervals, having been followed to new addresses if necessary, and asked additional questions to cover the intervening period. Any other persons that they had moved in with, or vice versa, were also interviewed. This continued for 15 months, and ended with a set of questions on taxes. In this way a highly detailed record was built up for each person and household for an entire calendar year. This design minimized the need for sample persons to recall the details of income and other characteristics for more than a few months and reduced the number of questions that had to be answered in each interview.

Because less time was required to update the basic information after the first interview, time was available in later interview waves to ask questions about other topics that were either stable enough not to require periodic updating--marital history and pension coverage, for example--or emerging issues of onetime interest, such as emergency energy assistance. This design enabled a set of core questions on income and other eligibility determinants to be developed well in advance, thereby, ensuring timely processing and rapid turnaround while leaving interview time to add questions on new policy issues on short notice.

Much of the work of the ISDP centered around four experimental field tests that were conducted to examine different concepts, procedures, questionnaires, recall periods, and the like. Two of the tests were restricted to a small number of geographic sites; the other two were nationwide. In the first nationwide test, the 1978 Research Panel, approximately 2,000 households were interviewed. Because of the relatively small number of interviews, controlled experimental comparisons of alternatives were not possible; however, the panel did demonstrate the feasibility of many new ideas and methods. It also laid a foundation for the largest and most complex test, the 1979 Research Panel. This panel consisted of a nationally representative sample of 8,200 households and provided a vehicle for feasibility tests and controlled experiments of alternative design features. Although used primarily for methodological purposes, it was sufficiently large to provide reliable national estimates of many characteristics of interest to analysts. (Public-use microdata files and documentation of the 1979 Research Panel are available through the National Technical Information Service. 1/) A more detailed discussion of the ISDP and its activities are provided in Ycas and Lininger (1981) and David (1983).

Because the ISDP was the predecessor to SIPP, it is not surprising that many characteristics of the ISDP are reflected in the SIPP design, including the survey design, content, and questionnaire format.

SIPP Design Features

SIPP started in October 1983 as an ongoing survey program of the Bureau of the Census with one sample panel of approximately 26,000 "designated" households in 174 primary sample units (PSU's) selected to represent the noninstitutional population of the United States. The actual sample size was somewhat smaller (about 21,000 households) because some of the selected households were unoccupied, demolished, converted for nonresidential use, or occupied by persons not eligible for interview, such as persons maintaining a usual residence elsewhere. The sample design is self-weighting; that is, each unit selected in the sample has the same probability of selection.

Each household is interviewed once every 4 months for 2 1/2 years to produce sufficient data for longitudinal analyses while providing a relatively short recall period for reporting monthly income. The reference period for the principal survey items is the 4 months preceding the interview. For example, in October, the reference period is June through September; when the household is interviewed again in February, it is October through January. This interviewing plan will result in eight interviews per household.

In February 1985 and then every January thereafter, a new, slightly smaller panel will be introduced. (Figure 1 illustrates the plan for implementing sample panels in 1984 through 1987.) This design will allow cross-sectional estimates to be produced from a combined sample of approximately 35,000 households.

Finally, to facilitate field operations, each sample panel is divided into four approximately equal subsamples, called rotation groups; one rotation group will be interviewed in a given month. Thus, one cycle or "wave" of interviewing takes 4 consecutive months. This design creates manageable interviewing and processing workloads each month instead of one large workload every 4 months; however, it results in each rotation group using a different reference period. Figure 2 provides an illustration of the relationship between waves, rotation groups, interview months, and reference periods.

To recap, the panels are an important feature of the SIPP design--a new panel is initiated each year There are also waves or interviews in each panel. Finally, there are rotation groups within each wave. Four rotation groups comprise a wave, each consisting of approximately one-fourth of the total sample. Since the reference period is the 4 months prior to the interview, one should notice that each rotation group has a different reference period.

The collection of data on a "staggered" basis produces 7 months of data under a 4-month reference period. This occurs because a full sample of cases during the wave is not available for each month of the reference period. Looking again at figure 2, January interviews obtain data for the period September through December; February interviews, for October through January; March interviews, for November through February; and April interviews, for December through March. By considering the use of individual wave files, the September data will only be available for the first rotation group; the October data, the first two rotation groups; the November data, the first three rotation groups; and so forth. Because of the design, however, matching individual wave data together will allow monthly analysis on the full sample. Although SIPP is essentially a monthly survey, the staggered design and the consequently staggered reference period only permit analysis on a full sample for one month of each wave.

SIPP Collection Procedures

Data collection operations are managed through the Census Bureau's 12 permanent regional offices. Interviewers assigned to these offices conduct one personal visit interview with each sampled household every 4 months. At the time of the interviewer's visit, each person 15 years old or older who is present is asked to provide information about himself/herself; a proxy respondent is asked to provide information for those who are not available. Telephone interviewing is permitted only to obtain missing information or to interview persons who will not or cannot participate otherwise.

The average length of the interview is about 30 minutes. An important design feature of SIPP is that all persons in a sampled household at the time of the first interview remain in the sample even if they move to a new address during the next 2 1/2 years. For cost and operational reasons, person-visit interviews are only conducted at new addresses that are in or within 100 miles of a SIPP primary sampling unit. After the first interview, the SIPP sample is a person-based sample, consisting of all individuals who were living in the sample unit at the time of the Wave 1 interview. Individuals aged 15 and over who subsequently share living quarters with the original sample people will also be interviewed in order to provide the overall economic context of the original sample persons. Changes in household composition caused by persons who join or leave the household after the first interview are also recorded. These individuals are interviewed as long as they reside with an original sample person.

Another important feature of SIPP is the identification numbering system. Each person will be assigned a unique fourteen-digit identification (ID) number at the time he/she enters the sample; an additional two-digits will be assigned if the person moves to a new address. A master list of identification numbers will be used by the regional offices to monitor the status of interviewing each month after Wave 1. The regional offices will be responsible for ensuring that there is a completed questionnaire (or reason for noninterview) for each number on or household, as well as income. Some of the major types of assets collected are dividend earnings, home equity, durable goods, and unsecured liabilities. This module, unlike the core data on assets, collects the value of the assets such as the market value of real estate, vehicles owned, and stocks held. This module will be administered twice in each panel, in waves one year apart.

Similarly, module of questions on tax-related information will be fielded twice in each panel. Information on filing status and taxes paid (income, property, and social security (FICA) taxes) will be obtained to allow the estimation of tax incidence, disposable income, and the simulation of tax policy alternatives. In addition to tax-related matters, an annual "round-up" module will also be administered to obtain wage and salary data from W-2 forms and estimates of annual self-employment and property income for each appropriate person.

Other fixed topical modules which will be administered in only one wave of the survey include: 1) marital history, 2) fertility, and 3) migration.

SIPP Content: Variable Topical Modules

In response to program planning and policy analysis data requirements, the final component of the SIPP content consists of modules of questions which will not be a recurring feature of each SIPP panel.

Brief descriptions of several of these variable topical modules will indicate the breadth of information collected. For example, in the third interview a module of questions on health care utilization and planning, and social services in-kind (health and training) was asked. In the fourth interview, a retirement and pension module contained questions on coverage and vested rights in retirement and pension plans. The data will help in the analysis of how net worth is related to retirement decisions and will allow a comparison of the Social Security system with private retirement plans.

Another topical module in the fourth interview collected characteristics of households that affect energy usage. These data will help provide a better estimate of income remaining after all housing needs are met and fulfill a need for information concerning energy usage. They will also provide information to allow the simulation and analysis of individuals and households qualifying for energy assistance programs.

In the fifth interview, a child care topical module has been developed to obtain information about child care arrangements, such as who provides the care, the number of hours of care per week, where the care is provided, and the cost of the care. These data will be useful because child care expenses are a major part of work-related expenses and are frequently deductible for program eligibility purposes.

In the same interview, questions on welfare history and child support will help determine the length of time persons receive aid from specific welfare programs, as well as provide information on child support agreements. The data from the welfare history questions will measure the extent to which persons and households have been dependent upon government transfer programs. Questions concerning child support will measure the degree to which the failure of the father to provide child care affects the likelihood of the mother and children's participation in government transfer programs.

A topical module on reasons for not working/reservation wage will include questions to ascertain why persons are not in the labor force. The data collected will aid the understanding of the conditions required for unemployed persons to accept a job.

Yet another topical module will contain questions about providing regular payments for the support of persons who are not members of the SIPP household and about expenses associated with a person's job. Those questions will help in obtaining a measure of the fixed financial obligations of persons, resulting in a more complete picture of their economic situation.

As can be seen, a wide variety of topics are covered under the aegis of the variable topical module concept. The breadth of these data in combination with the income and asset information ensure that SIPP will be a widely used and powerful data base serving many purposes.

SIPP Data Products

A number of publications and public-use data files will be generated from the information collected in SIPP. Both publications and data files are identified by whether they are cross-sectional or longitudinal. Two types of cross-sectional reports are planned by the Census Bureau: 1) a set of quarterly and annual reports that focussing on core information; and 2) a set of periodic or single-time reports using the detailed data from the topical modules.

The quarterly cross-sectional reports provide average values for a variety of labor force, income, and household composition measures based on monthly averages. The first quarterly report was issued in fall 1984 and contained income and labor force data referring to the third quarter of 1983. The annual reports will be similar in content, but will show values averaged across 12 months rather than 3 months. The periodic and single-time reports will use the detailed data from the topical modules to examine issues related to income and program participation. These reports may also focus solely on the material covered in a topical module such as work history or migration.

Plans for longitudinal data reports are under discussion. Six kinds of reports have been proposed for consideration (McMillen and Kasprzyk (1984)):

1) economic profile reports, presenting yearly aggregates of monthly data on individuals--reporting household and family information as characteristics of individuals;
2) comparative profile reports, presenting comparisons of yearly aggregates of monthly data on individuals;
3) transition reports, providing changes in

income and program participation status between two points in time;
4) multiple transition reports, providing patterns of labor force, income, and program participation activity, and the number of spells in a given state and their duration;
5) longitudinal family and unrelated individual reports, presenting the characteristics of longitudinal family units defined in SIPP (see McMillen and Herriot (1984)); and
6) special event reports, providing data related to an event, such as marriage, divorce, separation, the birth of a child, a return to school, or a move to a new address.

SIPP cross-sectional data files will be issued on a wave-by-wave basis. Each file will include person, family, and household information collected in the survey wave. Almost all data obtained on the questionnaire will be included on the files; certain summary income recodes will also be included. Data that might disclose the identity of a person will be excluded or recoded in accordance with standard Census Bureau confidentiality restrictions. Wave files will be edited, imputed, and weighted in a manner consistent with their use for cross-sectional analysis. A unique identification number will be included to allow users to merge two or more SIPP files. However, since the processing of wave files is independent, wave-to-wave data inconsistencies will occur and the user must be prepared to resolve them.

Plans for producing public-use files designed for longitudinal analysis are less well-defined at this time. The basic longitudinal file for SIPP is a calendaryear file of core income data; that is, the data items on the file will be monthly observations for 12 months of the calendar year for each person in the sample. Thus, the first of the series would be a calendar-year (CY) 1984 file of persons from the 1984 Panel.

SIPP Unit Nonresponse Rates

The first SIPP interviews were conducted in October 1983. At this time, cross sectional unit noninterview rates are available for the first three waves of SIPP. Unit noninterview rates (type A rate) provide a measure of the success/failure of the SIPP field work. While refusals are the largest part of the type A rate, it also includes "no one home" and "temporarily absent" households. Figure 3 provides an overall look at these rates. In Wave 1 (all rotation groups), the type A rate was 4.87 percent; in Wave 2, 3.73 percent; in Wave 3, 5.59 percent. These rates are an improvement on the rates experienced in the Income Survey Development Program.

SIPP-Related Research: Discussion of SIPP Research Issues

Interruptions in the funding support for SIPP in 1982 resulted in a cessation of the analysis associated with the pilot surveys of the development program. These surveys provided a large body of data to address a number of important methodological and substantive issues. Many of these issues were raised by Kasprzyk (1983) and discussed further in such forums as the Office of Management and Budget's (OMB) SIPP Advisory Committee, the Social Science Research Council's Subcommittee on the SIPP, the staff of the Committee on National Statistics, and the American Statistical Association's Census Advisory Committee. This year's meetings of the American Statistical Association are being used to bring the research community up-to-date on a variety of SIPP-related research issues. Topics, both methodological and substantive, are covered in four sessions organized under the auspices of the Social Statistics and Survey Research Methods Sections.

Finally, a SIPP working paper series has been established as a mechanism to provide timely and widespread access to information developed as part of the SIPP. Papers in the series will cover a broad range of topics including:
1) procedural information on the collection and processing of data;
2) survey methodology research; and
3) preliminary substantive results, such as the measurement of household composition change over time.

A substantial effort has been made by the SIPP staff to exchange ideas related to SIPP research with the research community. Several areas of research are described below.

SIPP Research: Accessing SIPP Data

Processing experience with data collected during the development program has shown that the complexity of the data, especially its longitudinal aspects, results in severe difficulties for the analysts. Indeed, the structure of the SIPP cross-sectional data files will seem very complicated to most users. The structure, described by Fink (1984), consists of a five-level hierarchy--sample unit, address or household, family, person, income types--with multiple record types in the fifth or lowest level. This structure, while chosen to provide maximum flexibility for cross-sectional data analysis and to simplify the merging of multiple waves of data into a longitudinal data base, does suggest that alternative ways of accessing SIPP data are necessary.

Acknowledging the data access difficulties, Census Bureau staff has been working with an OMB subcommittee comprised of representatives from various Federal agencies. This subcommittee has proposed the development of an "alternative" SIPP cross-sectional file. The proposed file has a rectangular structure and the individual is the basic unit of analysis. At this time it seems that the content of the rectangular file will not differ substantially from the more complicated file.

An internal Census Bureau committee investigated the need for and use of a data base management system for SIPP within the Census Bureau's current processing environment. The study concluded that Census Bureau options in the statistical data base field were quite limited and that, of the data base management systems (DBMS) the list representing all the persons assigned for interview in a month. The list will be updated regularly to account for persons who are added or deleted from the sample.

The ID is intended to provide a means of linking information about an individual across time and uniquely identifying which household each person is a member of at any point in the panel. Through the ID system, we expect to link data from all persons ever associated with a sample unit throughout the 2 1/2-year duration of a panel. This will facilitate the construction of household income estimates based on the actual composition of households during the measurement period. More information about the construction and use of the ID number can be found in Nelson, McMillen, and Kasprzyk (1984) and Jean and McArthur (1984).

SIPP Content: Control Card

The control card is used to obtain and maintain information on the basic characteristics associated with households and persons and to record information for operational control purposes. Characteristics recorded on the control card by the interviewer include the age, race, ethnic origin, sex, marital status, and educational level of each member of the household, as well as information on the housing unit and the relationship of the householder to other members. A household respondent provides this information, which is updated at each interview. The control card is also used to keep track of when and why persons enter and leave the household, thereby providing enough information to automatically create monthly household and family groups. There is also space to record information that will improve our ability to follow persons who move during the survey. In addition, after each visit, data on employment, income, and other information is transcribed from the core questionnaire to the control card so the data can be used in the next interview.

SIPP Content: Core Data

The content of SIPP was developed around a "core" of labor force and income questions designed to measure the economic situation of persons in the United States. These questions expand the data currently available on the distribution of cash and noncash income and are repeated at each interviewing wave. SIPP core data build an income profile of each person aged 15 and over in a sample household. The profile is developed by determining the labor force participation status of each person in the sample and asking specific questions about the types of income received, including transfer payments and noncash benefits from various programs for each month of the reference period. A few questions on private health insurance coverage are also included in the core.

Persons employed at anytime during the 4-month reference period are asked to report on jobs held or businesses owned, number of hours and weeks worked, hourly rate of pay, amount of earnings received, and weeks without a job or business.

In addition to questions about labor force activity and the earnings from a job, self-employment, or farm, the core includes questions related to nearly 50 other types of income. Questions about common income types are specifically asked while the receipt of less common income types are elicited through general probing questions. Questions are asked about the receipt of government transfer payments from retirement, disability, unemployment benefits, and welfare programs. Information on the receipt of noncash benefits from programs such as Medicare and Medicaid is also obtained. Other income questions in the core relate to private transfers such as pensions from employers, alimony, and child support. For certain income types, such as food stamps and AFDC, questions are included which help to identify the household members covered by the payment, thus allowing the proper construction of program analysis units. Finally, the core data also include questions on the ownership of assets which produce income, such as savings accounts, money market accounts, NOW accounts, stocks, mutual fund shares, and rental property. The amounts of income received from these income producing assets are also obtained, as well as indications of joint holdings and estimates of account balances if the amount of interest is not known.

SIPP Content: Fixed Topical Modules

The core data provide information on a continuing basis about levels of economic well-being and changes in these levels over time. These data, while extremely detailed, allow analyses of well-being which only account for income and demographic variables. The SIPP has been designed to provide a broader context for analysis by adding questions on a variety of topics not covered in the core section. These questions are labelled "fixed topical modules" and are assigned to particular interviewing waves of the survey. If more than one observation is needed, questions on one wave may be repeated on a later wave.

The administration of these modules of questions is made possible by the fact that less time is required to update the core information collected in the first interview. Also, the topics covered in these modules do not require repeated measurement at each interview and, therefore, may use a reference period longer than the period used for the information obtained in the core. For example, the third SIPP interview questionnaire collects information on health and disability, and education and work history. The former are obtained because they are among the major factors affecting a person's ability to work, his/her earnings, sources of income, and participation in public programs. The latter provide data to help understand a person's economic situation in relationship to his educational and occupational background. The fourth interview contains topical data on assets and liabilities, retirement and pension coverage, and housing conditions/energy usage. The collection of assets and liabilities data allows the study of economic well-being beyond that which can be observed through the study of income alone. Participation in many Federal programs is contingent upon assets held by the individual of the SIPP to correct and verify a respondent's SSN.

Having established the link for matching activities, work is now proceeding on identifying content and availability of administrative record systems for use in: a) data augmentation for research and estimates; and b) survey data evaluation. In the former, a working group is developing 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. Information concerning this study can be found in Haber, Ryscavage, Sater, and Valdisera (1984).

Another area of research with respect to administrative record systems is the development of validation studies of items common to both the survey and administrative records. The goal of the project is the improved understanding of the quality of the SIPP data and, ultimately, the development of quantitative estimates of response and nonresponse errors for the purposes of adjusting survey data or modifying survey procedures to obtain better quality data. The first aspect of this work is the development of a research plan identifying activities required to evaluate SIPP data using administrative records. The second aspect is a demonstration and feasibility study to examine response and nonresponse errors from Waves 1 and 2 of the SIPP using administrative record systems for the major transfer programs from a limited number of states.

SIPP Research: Panel Surveys as a Source of Migration Data

SIPP data can be used to address a wide variety of migration topics in two ways: 1) by traditional--cross-sectional--analyses, these data serve to further understanding of how geographical mobility leads to adjustments in labor markets, housing markets, etc.; and 2) the survey's longitudinal design provides a natural source of geographical mobility data because individuals are followed should they move to a new place of residence. The first stage of this work is to review analyses of migration from previous panel surveys and then assess how SIPP can further our understanding of geographical mobility processes. Some results from this work are described by Dahmann (1984).

SIPP Research: Wave-to-Wave Changes in Income and Program Receipt

Analysts with an interest in the quality of the data obtained in the 1979 ISDP Panel--and the subsequent SIPP--have expressed concern that respondent reports of income receipt may be flawed. There appears to be a tendency for reported program turnover to occur between waves more often than within waves--that is, in the pilot surveys of the development program, between months 3 & 4 rather than the other 4 consecutive pairs of months. It is assumed that this probably represents response error arising from imperfect recall although other factors (for example, mismatching in linking the data files) might also account for this effect. Some analyses of this phenomenon using five waves of data from the 1979 Panel and examining the extent to which differences are related to respondent status patterns across survey waves are presented in Moore and Kasprzyk (1984). Additional work using SIPP data is planned as soon as data are available.

SIPP Research: Longitudinal Feedback and Reconciliation System

Because of its design, SIPP has a potential for missing and inconsistent data problems from wave to wave. The issue of concern is the development of appropriate forms and procedures to identify and correct longitudinal data problems during the collection rather than the processing phase. An automated income and work experience profile to identify potential cross-wave edit failures and data problems could help in the development of SIPP longitudinal data products. This profile would contain responses on labor force activity and amounts of income and program benefits received during the previous calendar year. It would be reviewed for accuracy by each respondent at the conclusion of the calendar year. A system such as this should clarify apparent inconsistent responses by using previously reported amounts to identify and reconcile cross-wave inconsistencies.

Prior to designing and implementing a reconciliation system to smooth transition data during a calendar year, a preliminary review of a subsample of questionnaires for the first two waves of SIPP will be conducted.

SIPP Research: Sampling for Special Populations

After the SIPP becomes established Federal program agencies may be interested in adding sample cases for specific subpopulations of interest to policy analyses, such as the high and low income groups, Blacks and Hispanics, and the aged and disabled. A multi-divisional work group is discussing methods for oversampling special populations. In particular, a variety of subsampling (screening) proposals will be analyzed.

The statistical issue under consideration is the reliability of estimates when different subsampling schemes are introduced. During the last several months, subsampling characteristics based on income and demographic variables have been identified, and estimates of reliability have been obtained for different subsampling rates and different characteristics. A summary of results is now in preparation.

SIPP Research: Sampling Error Estimation

Applications of SIPP data are expected for a wide variety of analyses--microsimulation modeling, multivariate analysis, and simple tabulations and cross-tabulations. Some topics under study in the Statistical Research Division are: 1) an investigation of currently available computer software which provide general procedures for computing sampling error estimates for a complex survey designs; and
2) an assessment of Census Bureau procedures for computing and estimating the median and its variance.

available at the Census Bureau, the Scientific Information Retrieval (SIR) DBMS seemed to be the most suitable for SIPP. As a result of this study, SIPP data are now being structured in the SIR data base format by the Census Bureau's Systems Support Division.

Simultaneously, along with the SIR data base construction, consideration is being given to the development of a software system to accompany SIPP data files and allow the relatively easy generation of extract files of records focusing on changes, duration, and spells of a particular status. These extract files would serve as an intermediate product to be used as input into the more widely available statistical software packages.

Finally, to improve the understanding of and access to SIPP data, a data product and delivery program is being developed. This program would include introductory SIPP workshops and user workshops, newsletters, and guides to increase the understanding of SIPP data products.

SIPP Research: Longitudinal Concepts-Household and Family Definition

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 traditional cross-sectional household and family constructs. McMillen and Herriot (1984) discuss this problem in detail by examining cross-sectional concepts and their deficiencies, examining and evaluating several longitudinal concepts, proposing a concept for SIPP, and then illustrating how the concept would be used in calculating aggregate household/family characteristics and in tabulating the number of households, household types, and their characteristics.

SIPP Research: Longitudinal Estimation

Weighting the longitudinal sample, especially for analytic units other than the individual, is a difficult area requiring a continuing statistician-analyst dialogue. Detailed longitudinal weighting procedures will not be developed until closure is reached on definitions for longitudinal analysis units. Despite the lack of closure, 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 longitudinal person estimation. Preliminary thoughts on this topic are described by Judkins, Hubble, Dorsch, McMillen, and Ernst (1984).

The topic of longitudinal household (family or recipient unit) estimation is also under study. The work includes the development of an explicit statement of the longitudinal estimation problem, the survey universe under several definitions of longitudinal household, and the principal estimation issues which require study. Preliminary thoughts on this topic are discussed in Ernst, Hubble, and Judkins (1984).

SIPP Research: Longitudinal Imputation

The varieties of types of nonresponse--unit nonresponse defined as nonresponse to all waves of the survey, wave nonresponse defined as nonresponse to a particular wave interview, and item nonresponse defined as nonresponse to a particular item--and their patterns pose numerous difficulties for designing appropriate strategies for nonresponse compensation. At this time several aspects of the problem are being developed. The Statistical Research Division at the Census Bureau is studying patterns of item "missingness"; frequency of state-to-state transitions; and evaluating proposed imputation strategies, both the model-based and hot-deck types, for labor force, wages, and salary items on the 1979 ISDP Panel. Some preliminary ideas on the treatment of multiwave item nonresponse in the SIPP are discussed in Samuhel and Huggins (1984).

In addition, two other aspects of the problem are also under development:

1) To study item nonresponse for property and program income--in particular, examining levels and patterns of item nonresponse and developing and comparing methods to treat nonresponse in these subject areas using more than one wave of data. This work is similar to the previously discussed work on labor force, wages, and salaries, but in a different subject-matter area.
2) To study general strategies of handling missing data in panel surveys, including: (a) a discussion and analysis of issues pertaining to weighting versus imputing for attrition cases; (b) a discussion of the treatment (weighting or imputation) of the so-called "non-nested" missing data cases; (c) an empirical examination and comparison of weighting and imputation using data from the 1979 ISDP Panel; and (d) a discussion based on empirical results of when to choose one strategy over the other.

SIPP Research: Use of Combined Survey and Administrative Data

During its development period, SIPP had been viewed as an integrated data system, combining survey data with administrative record data. Because of this emphasis, an internal Census Bureau committee was formed to assess and make recommendations regarding the potential uses of administrative records, the development of demonstration pilot studies, and the special confidentiality or privacy issues involved in the use of administrative records. The committee developed a proposal, later implemented, to electronically validate reported social security numbers (SSN), to manually search for SSN's not reported correctly, and to use the panel aspect

SIPP Research: The American Statistical Association-Census Bureau Research Fellow Program

Recognizing that SIPP research cannot be confined solely within the Census Bureau, SIPP research planning has encouraged the use of development program data to increase the understanding of the SIPP data and the inherent difficulties working with this new data base. One aspect of this activity has been the expansion of the ASA-Census research fellow program to identify explicitly SIPP-related research activities. As a result, two research fellows have been chosen for the 1984-1985 academic year: Harold Watts (Columbia University) and Constance Citro (National Academy of Sciences and Mathematica Policy Research). Dr. Watts intends to use the 1979 Panel data to understand changes in living arrangements and how long they last. He is interested in short term gross flow patterns of change in household status and will characterize the changes in the household status of the individual from one wave to another. Dr. Citro intends to simulate alternative definitions of household continuity. She will develop tables of annual household and family income under alternative definitions using the data from the development program's 1979 Panel and will help interpret the meaning and usefulness of the alternative measures.

Summary

This paper has reviewed the development and current activities of the Survey of Income and Program Participation. Survey design features, selected field procedures, content of the survey, and data products have been described. As a number of methodological issues remain unresolved, selected research activities have been described. While covering many aspects of the SIPP, this paper was not meant to be comprehensive. The scope of topics discussed, however, illustrates the nature, breadth, and opportunities of the Survey of Income and Program Participation.

1/ To request tapes and documentation describing the history of the 1979 Research Panel, sample design, survey content, estimation procedures, and data collection and processing procedures, write to: Department of Commerce, National Technical Information Service, 5385 Port Royal Road, Springfield, Virginia 22161 or call (703) 487-4807.

BIBLIOGRAPHY

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

David, M. (ed), (1983). Technical, Conceptual, and Administrative Lessons of the Income Survey Development Program. New York, Social Science Research Council.

Ernst, L., Hubble, D., and Judkins, D. (1984). Longitudinal Family and Household Estimation in SIPP. Proceedings of the Survey Research Methods Section, American Statistical Association.

Fink, Barry (1984). 1984 SIPP File Structure, Memorandum for Members of the SIPP Committee for Data Products and Processing.

Haber, S., Ryscavage, P, Sater D., and Valdisera, V. (1984). Matching Economic Data to the Survey of Income and Program Participation: a pilot study. Proceedings of the Social Statistics Section, American Statistical Association.

Jean, A., and McArthur, E. (1984). Some Data Collection Issues for Panel Surveys with Application to the Survey of Income and Program Participation. Proceedings of the Survey Research Methods Section, American Statistical Association.

Judkins, D., Hubble, D., Dorsch, J., McMillen, D. and Ernst, L. (1984). Weighting of Persons for SIPP Longitudinal Tabulations. Proceedings of the Survey Research Methods Section, American Statistical Association.

Kasprzyk, D. (1983). Some Research Issues for the Survey of Income and Program Participation. Proceedings of the Survey Research Methods Section, American Statistical Association.

McMillen, D., and Herriot, R.A. (1984). "Toward a Longitudinal Definition of Households." SIPP Working Paper Series No. 8402. U.S. Bureau of the Census, Washington, D.C. 1984.

McMillen, D., and Kasprzyk, D. (1984). Memorandum for the Record, "Potential SIPP Longitudinal Data Products," May 1984.

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

Nelson, D., McMillen, D., and Kasprzyk, D. (1984). "An Overview of the Survey of Income and Program Participation." SIPP Working Paper Series No. 8401. U.S. Bureau of the Census, Washington, D.C. 1984.

Samuhel, M. and Huggins, V. (1984). Longitudinal Item Imputation in a Complex Survey. Proceedings of the Survey Research Methods Section, American Statistical Association.

Ycas, M., and Lininger, C. (1981). The Income Survey Development Program: Design Features and Initial Findings. Social Security Bulletin, Vol. 44, No. 11, November 1981.

Figure 1. SIPP Interview Schedule

SIPP INTERVIEW SCHEDULE					
MONTH YEAR	1984	1985	1986	1987	DATA MONTHS
OCT 83	I				JUNE-SEPT
NOV	I				JUL-OCT
DEC	I				AUG-NOV
JAN 84	I				SEPT-DEC
FEB	CU				OCT-JAN, 1984
MAR	CU				NOV-FEB
APR	CU				DEC-MAR
MAY	CU P2				JAN-APR
JUN	CU P2				FEB-MAY
JUL	CU P2				MAR-JUNE
AUG	CU P2				APR-JULY
SEP	CU W1				MAY-AUG
OCT	CU W1				JUNE-SEPT
NOV	CU W1				JULY-OCT
DEC	CU W1				AUG-NOV
JAN 85	CU VT	I			SEPT-DEC
FEB	CU VT	I			OCT-JAN, 1985
MAR	CU VT	I			NOV-FEB
APR	CU VT	I			DEC-MAR
MAY	CU R1	I			JAN-APR
JUN	CU R1	CU			FEB-MAY
JUL	CU R1	CU			MAR-JULY
AUG	CU R1	CU			APR-JULY
SEP	CU W2	CU W1			MAY-AUG
OCT	CU W2	CU W1			JUNE-SEP
NOV	CU W2	CU W1			JULY-OCT
DEC	CU W2	CU W1			AUG-NOV
JAN 86	CU P1	CU P1			SEPT-OCT
FEB	CU P1	CU P1	I		OCT-JAN, 1986
MAR	CU P1	CU P1	I		NOV-FEB
APR	CU P1	CU P1	I		DEC-MAR
MAY	CU R2	CU R1	CU TM		JAN-APR
JUN	CU R2	CU R1	CU TM		FEB-MAY
JUL	CU R2	CU R1	CU TM		MAR-JUNE
AUG	CU R2	CU R1	CU TM		APR-JULY
SEP		CU W2	CU W1		MAY-AUG
OCT		CU W2	CU W1		JUNE-SEP
NOV		CU W2	CU W1		JULY-OCT
DEC		CU W2	CU W1		AUG-NOV
JAN 87		CU P2	CU P2	I	SEPT-DEC
FEB		CU P2	CU P2	I	OCT-JAN, 1987
MAR		CU P2	CU P2	I	NOV-FEB
APR		CU P2	CU P2	I	DEC-MAR
MAY		CU R2	CU R1	CU TM	JAN-APR
MAY		CU R2	CU R1	CU TM	JAN-APR
MAY		CU R2	CU R1	CU TM	JAN-APR
JUN		CU R2	CU R1	CU TM	FEB-MAY
JUL		CU R2	CU R1	CU TM	MAR-JUNE
AUG		CU R2	CU R1	CU TM	APR-JULY
SEP			CU W2	CU W1	MAY-AUG
OCT			CU W2	CU W1	JUN-SEPT
NOV			CU W2	CU W1	JULY-OCT
DEC			CU W2	CU W1	AUG-NOV

I - Initial interview

CU - Update

P1 - Personal history (1) (Marital history, fertility, migration)

P2 - Personal history (2) (Health and disability, work history, education history)

W1 - Assets & Liabilities

W2 - Assets & Liabilities

VT - Variable topical module-child care, arrangements duration of welfare

TM - Unassigned topical module

R1 - First annual roundup (Annual roundup - annual income, taxes, educational enrollment)

R2 - Second annual roundup (Annual roundup - annual income, taxes, educational enrollment)

Figure 2. SIPP Interview Months and Reference Periods

Wave	Rota-tion Group	Inter-view Month	Reference Periods																			
							1st Qtr Date			2nd Qtr Date			3rd Qtr Date			4th Qtr Date						
1	1	JAN	SEP	OCT	NOV	DEC																
	2	FEB		OCT	NOV	DEC	JAN															
	3	MAR			NOV	DEC	JAN	FEB														
	4	APR				DEC	JAN	FEB	MAR													
2	1	MAY					JAN	FEB	MAR	APR												
	2	JUN						FEB	MAR	APR	MAY											
	3	JUL							MAR	APR	MAY	JUN										
	4	AUG								APR	MAY	JUN	JUL									
3	1	SEP									MAY	JUN	JUL	AUG								
	2	OCT										JUN	JUL	AUG	SEP							
	3	NOV											JUL	AUG	SEP	OCT						
	4	DEC												AUG	SEP	OCT	NOV					
4	1	JAN													SEP	OCT	NOV	DEC				
	2	FEB														OCT	NOV	DEC	JAN			
	3	MAR															NOV	DEC	JAN	FEB		
	4	APR																DEC	JAN	FEB	MAR	
	1	MAY																	JAN	FEB	MAR	APR

Figure 3. SIPP Noninterview Data: Waves 1, 2, and 3

	Wave 1		Wave 2 1/		Wave 3 2/	
	Number	Percent	Number	Percent	Number	Percent
Households Assigned	26024	100.00	15358	100.00	20909	100.00
Type B Noninterview	3362	12.91	0	0.00	0	0.00
Type C Noninterview	1765	6.78	97	0.63	132	0.63
Eligible Households	20897	80.29	15261	99.36	20777	99.37
Interviewed Households	19878	95.12*	14532	95.22*	19190	92.36*
Noninterviewed Householdx	1019	4.87*	729	4.77*	1587	7.64*
Type A Noninterviews	1019	4.87*	570	3.73*	1161	5.59*
Refused	771	3.68*	468	3.06*	958	4.61*
Other	248	1.19*	102	.67*	203	.98*
Type D Noninterviews	0	0.00*	159	1.04*	426	2.05*

Type A - Noninterview
Refused
Other (no one home, absent, unable to locate)

Type B - Noninterview
Vacant
Occupied by persons with URE
Unfit or to be demolished
Under construction, not ready
Converted to temporary business or storage
Unoccupied tent site or trailer site
Permit granted, construction not started
Other

Type C Noninterview (Wave 1)
Demolished
House or trailer moved
Converted to permanent business or storage
Merged
Condemned
Other

Type Noninterview (Waves 2-3)
Entire household out-of-scope (For example: entire household deceased, institutionalized, etc.)

Type D - Noninterview
Moved, address unknown
Moved within country beyond limit

*/ Percentage of eligible households.
1/ Note, Wave 1, Type A noninterview households were not reassigned in Wave 2.
2/ Wave 3 data are preliminary. Note: That households assigned in Wave 3 include Wave 2 Type A noninterview households.

THE GERMAN SOCIO - ECONOMIC PANEL

Ute Hanefeld, Deutsches Institut für Wirtschaftsforschung, Federal Republic of Germany

Introduction

The German Socio - Economic Panel was started with its first wave in the spring of 1984. It will provide representative longitudinal data on income, transfer payments, labor market experience, changing family composition, and housing for individuals, families and households. In addition, it will allow representative cross-sectional analyses. It is the first study of this kind in the Federal Republic of Germany.

The panel is designed for a variety of analyses which could not be conducted for West Germany without this data base. Primarily, the dynamics of individual behavior and personal characteristics are of interest. For example, income changes of individuals, the causes of these changes and their consequences will be analyzed. Another area are changes in labor market participation, particularly for women. Patterns of behavior and their interdependencies with the living conditions of the household and the behavior of the other household members will be studied.

The sample is representative of the entire population of West Germany, including foreigners and, at least partially, institutionalized persons. All household members 16 years and older are interviewed annually in 6 000 households. Each initial sample member will be followed within the area of the Federal Republic of Germany in the future. At present, data collection for at least five years is planned. If the study is successful, it is intended to add further waves.

The study is conducted by the Special Collaborative Programme 3 'Microanalytical Foundations of Social Policy' (Sonderforschungsbereich 3 'Mikroanalytische Grundlagen der Gesellschaftspolitik') at the Universities of Frankfurt and Mannheim, in collaboration with the 'German Institute for Economic Research' (Deutsches Institut für Wirtschaftsforschung, DIW) in Berlin. The fieldwork is done by Infratest Sozialforschung, a survey institute in München. The project is funded by the German Research Society (Deutsche Forschungsgemeinschaft, DFG) in Bonn.

Aims of the study

Up to now, in the Federal Republic of Germany only a few panel studies have been conducted in social sciences. These studies focus on individuals and in general do not take the household into account. In addition, they often have very small sample sizes and concentrate on special groups of the population and on special topics. Some examples are: panels that investigate the transition of youth from education to the labor market (1), fertility studies (2) and one panel between 1978 and 1982 concerning unemployment(3). Panel studies concentrating on questions about income, transfer payments or labor market experience do not exist.

With the Socio - Economic Panel, a broader purpose is intended:

- It is designed to be representative of the entire population of West Germany.
- It will allow longitudinal analyses on the level of individuals, families and households.
- It is a multi-purpose survey covering a broad range of aspects concerning living conditions.

Usually surveys for the Federal Republic of Germany are only representative of the German population living in private households (4). Because of language and sample selection problems, foreigners are not sampled, a group to which 7.4 % of the population belonged in 1983. Further, the institutionalized population (5) is not included. Even the German Federal Statistical Office (Statistisches Bundesamt) does not take into account these two groups in most surveys except the annual microcensus which is comparable to the Current Population Survey of the U.S.

Because of the importance of guestworkers in many aspects of today's economy, this group has been included in the Socio - Economic Panel and oversampled to allow separate analyses. The institutionalized population with 1.5 million persons in 1970 (6), is not negligible either. In the official definition, it consists to a large part of elderly persons living in old people's homes and of employees living in hostels. People who might be more difficult to be interviewed because they are ill, living in prisons or barracks, only amount for less than half of the whole group. Additionally, the exclusion of the institutionalized population from the panel would lead to substantial panel mortality if persons were not followed when they move from a household to an institution. Therefore, an attempt has been made to include this population in the panel, too.

The study is designed to provide longitudinal data for individuals as well as for families (7) and households. Many problems refer to individuals, e.g. labor market experiences or earnings. However, it is also important to know the family and household background of the person. For other topics, like poverty or housing, information on the household level is needed. Analyses of transfer programs require data for families and households.

Cross-sectional analyses have shown that during the last years the size of households decreased and the number of households increased. But to a large extent it is unknown how this process works and which are the causes. With the Socio - Economic Panel it will be possible to investigate changes in the household composition and their causes and circumstances. At the same time, the influence of changing household composition on living conditions of the household members is of great interest. For example, one might ask for the impact of divorces or death on family income or on the labor market participation of the household members.

For the Federal Republik of Germany little is known about income changes on the level of households and individuals. It is planned to analyze the frequency and extent of such income changes, their causes and how people cope with such situations. Investigations of the poverty population will be of special interest. Which factors are responsible for the entry into low income groups? How much fluctuation does exist within the poverty population? Who is able to leave the poverty status and by which means? In connection with such questions, it will be possible to analyze transfer program participation, and whether different transfer programs reach their target groups and their goals.

During the last ten years unemployment became a severe problem in the Federal Republic of Germany. Changing demographic structures increase the difficulties on the labor market because at the moment large cohorts enter the labor force. This situation intensifies the necessity of longitudinal analyses, especially of occupational careers and mobility.

Concerning labor supply, information about the patterns of behavior of women and youth are important. With the panel study, it will be possible to analyze female labor force participation, interrupts in labor supply, for how long, and what the possibilities and circumstances are to reenter the labor market.

As mentioned above, there exist already some panel studies in the Federal Republic of Germany concentrating on the transition of youth from education to the labor market. Because these studies observe more cases of the particular subgroups than the Socio-Economic Panel, analyses based on these data can be more detailed. But as these studies concentrate on persons, and because some of them only follow selected cohorts, it is the task of the Socio-Economic Panel to supplement household related information and to check whether the results concerning special cohorts can be generalized.

Another important area of the Socio-Economic Panel will be the analysis of job mobility. Until now, only little is known about who is changing jobs and for what reasons, how people find a new job, what career patterns do exist and what are the consequences of job changes for other living conditions.

The fourth main area of the panel concerns housing and residential mobility. Although, since the Second World War, government has supported the housing market with diverse programs, housing is still an important theme in political debates. Cross-sectional analyses have shown that there is up to now a remarkable unbalance in the distribution of quality and quantity of housing. Growing aspirations and an increasing number of households are discussed as causes for this situation.

With the Socio-Economic Panel it will be possible to study the development on the housing market and the efficiency of transfer programs related to housing. For example, it is planned to analyze which households move, what the causes are for residential mobility, who is changing from the renter status to owner status and vice versa.

Beside the main topics of the study, changing family composition, income, transfer payments, labor market experience and housing, there are questions about education, health and attitudes.

It is not possible to mention all the planned analyses here. But as can be seen by the broad range of topics, the Socio-Economic Panel is designed as a multipurpose survey. The main interest of the study is to investigate not only the changes within the particular areas, but also the interdependencies among them. On the other hand this implies, that analyses in a single area can not be as detailed as in special studies conducted for that purpose. As an example, information about unemployment will not be as detailed as in a specialized panel, but the Socio-Economic Panel will show such relationships as the consequences of unemployment for the labor market participation of other household members, for family income or for residential mobility.

<u>The sample</u>

As explained above, the population of the Socio-Economic Panel is defined as the entire population of the Federal Republic of Germany including guestworkers and institutionalized persons. However, general problems arose with regard to the sample selection. Initially, it was intended to built up a new sample frame based on the census which was planned for 1983 (8). Unfortunately, this census was cancelled because of the growing discussion about privacy issues. In summary, since other approaches proved to be infeasible, there was no sample frame available covering the entire population. Even the existing sample frame for the microcensus could not be used because the samples were not ready, and the work that still needed to be done could not be completed because of the political controversy over invasion of privacy. Besides, this sample frame probably is biased since it is still based on the 1970 census.

Therefore, two subsamples were constructed and for each a special sample was drawn:

Sample A:	Persons in households with a head who is not Spanish, Italian, Greek, Turkish or Jugoslavian.
Sample B:	Persons in households with a Spanish, Italian, Greek, Turkish or Jugoslavian head.

For sample A, an existing sample frame used by the German survey institutes (9) was modified. It is an area sample based on voting districts of the last election. Because guestworkers are not allowed to vote, they are not represented in this frame. For the 4,500 households of sample A, 584 sample points were selected systematically with a random start. In each sample point the interviewer got a starting address. Starting from this address, the interviewer had to note the following 84 addresses. He or she then had to recruit every seventh household for the panel, but the last two addresses were only used if not at least eight interviews were obtained from the first ten households.

Sample B includes the main groups of guestworkers. It has a sample size of 1,400 households: 400 Turkish, 300 Italian, 300 Jugoslavian, 200 Greek and 200 Spanish households. For this population there are regional administrative registers. An area sample was drawn separately for each nationality. The sample points again were selected, by systematic choice with a random start. Altogether, 240 sample points were chosen and the administrations of the corresponding counties were asked to select 20 addresses at random. At least 7 addresses per sample point were passed to the interviewers to be recruited as panel households.

For the pretest - done in autumn 1983 with 200 households - a separate sample for the institutionalized population was tested. However, since no sample frame is available and data are insufficient, for the moment it is only possible to construct a quota sample. The results of the pretest were ambivalent. There were no special problems getting an interview from people living in old people's homes, hostels, dormitories etc. But big problems arose when interviewing people in barracks, sanatoria etc. There, in most cases, the interviewer was not even able to obtain permission to select a respondent and contact him.

It might be possible to draw a sample for the whole institutionalized population, if separate questionnaires - which take into account the special problems of this group - were designed and if special efforts were made to get the permission of the administration of the institutions. Since the subsample would have had a size of only 180 persons in the main waves, the necessary additional costs and work did not appear to be justified in view of the expected result. Therefore, the decision was made to include that part of the institutionalized population, where there were no extraordinary difficulties, in sample A or sample B, respectively. For the remaining part, it was decided that each sample member who moves into such an institution will also be followed. With this concept, it is hoped to cover the entire institutionalized population in the course of the panel.

Interviewing procedure

In many studies, only the head of the household is interviewed, but as the Panel Study of Income Dynamics has shown, the head changes quite frequently (10). For a panel, this would mean that different respondents answer the questionnaire in different waves. In addition, there is the risk that the head is not well enough informed about characteristics of other household members. Besides, information on attitudes is required. But, it is not feasible to ask the head of the household about attitudes of other household members , since subjective variables can only be measured appropriately by self-interviews. If only heads of households were interviewed, information would be biased. On the other hand, if the respondent were selected randomly in the household, the risk is even higher that he or she is not enough informed about household related questions. Therefore, it was decided to interview all household members 16 years and older in the Socio -Economic Panel.

Standardized fixed format questionnaires are used for the data collection. There are bilingual questionnaires for the foreign subsample. Each of the five nationalities gets a questionnaire in their mother tongue and under each question the German version is also printed. This form was chosen in case there were different nationalities in a household and to assist the German interviewer in keeping control over the interviewing process.

As none of the survey institutes has a permanent foreign language fieldstaff and because it is expensive and risky to build a new staff which would be lacking in interview experience, an unusual form of interviewing was tried in the pretest. Each interviewer in the foreign subsample had to find a foreigner of that nationality to accompany him or her. This person was trained very briefly by the German interviewer to help in making the first contact and, if necessary, in doing the interview. The pretest showed that, in general, this concept works well. However, in cases when the respondents knew German well, these accompanying persons sometimes were disturbing. Therefore, the rule was changed, so that the interviewer could use such a person, but is not required to do so.

To train the interviewers for the first wave, each got an interviewing package that included a test with several questions particularily concerning definitions of the population and the use of the coversheet. For the second wave, this test will focus mainly on questions dealing with how to follow respondents and whom to interview. In addition, the interviewers had to conduct one test interview of a sample household. The test and the test interview were controlled by the survey institute, and if necessary, the interviewer was retrained before he or she was allowed to do further interviews. This concept worked very well, although it should be mentioned that many interviewers did not pass this training, either because they did not want a work load considerably above that of usual surveys, or they made too many mistakes. This procedure resulted in 168 interviewers being exchanged who did not qualify. Finally, a total of 622 interviewers was engaged in the fieldwork.

The questionnaires

Three different questionnaires were used for the first wave of the Socio - Economic Panel. The coversheet is the main instrument to control the panel systematically. The interviewer has to fill it out for each sampled household. In the first wave, the name of the household had to be noted, and additionally for each address whether the contact was successful or the cause of a drop-out or a refusal. In a second part of the coversheet, the interviewer had to note all household members 16 years and older by name, gender, year of birth and relation to the head of the household. If a household member droped out or refused, the interviewer also had to note the special reason. In addition, there were some questions about the housing environment which provide supplementary information for the analyses of refusals. With this coversheet each sample member got an identification number.

Beginning with the second wave, each coversheet will be printed with the household address, and the identification number, name, gender and year of birth of all household members as of the last data collection. If the whole household has moved, the interviewer has to note the new address. He/she has to indicate whether new members belong to the household and, if this is the case, their main characteristics. If household members leave the household, the interviewer has to find out their new address and send it to the institute. For households that moved or split, the institute fills out a special coversheet and gives it back to an interviewer living in that area.

The identification number, the first name, gender and year of birth are noted on the other questionnaires as well and used, in combination with the information of the coversheet, to ensure that the respondents remain the same from one wave to the next.

The second questionnaire, called the household interview, is administered in an oral interview to the household head or a household member who is well informed about the whole household. It takes about 15 minutes. This questionnaire contains questions which do not need to be directed to every household member, either because the same information would be collected several times, such as information about the children, housing and household-related transfer payments, or because there is, for some variables, the risk that it is not clear whether the answer is household or person-related. An example of the latter are questions about property.

The third questionnaire is the personal interview. Because in large households it might be a heavy burden to conduct personal interviews with every household member 16 years and older, which might reduce the response rates, in sample A it is up to the interviewer to decide when to give an oral interview and when to have respondents fill out their own questionnaires. But, with some exceptions, the interviewer is to be present when the respondents answer the questionnaire, to motivate and help them if necessary. Because there are some foreigners who have difficulties in writing and reading, in Sample B the interviewers are instructed to conduct oral interviews.

The personal questionnaire takes about 30 minutes to complete. It contains questions about earnings, person-related transfer payments, labor market experience, education, health, time use, attitudes and demography. In the first two waves there are also biographical questions, especially retrospective questions about labor market participation and marital status. In each wave the questions on income, transfer payments and labor market participation cover the previous calendar year to record all changes. A detailed description is only collected for the current job at the time of the interview; this means that there is no information on any other jobs between two waves.

Starting with the second wave, there is a supplementary questionnaire to the personal interview. It is addressed to new sample members, being interviewed for the first time, persons who either were 15 years during the last wave, or persons who live together with initial sample members. The suppementary questionnaire contains the most important questions which were asked in previous waves and not repeated in the current wave. It includes, for example, the questions on the level of schooling and vocational training and the retrospective biographical questions.

Methods to maintain the panel

Although in every wave respondents of the Socio - Economic Panel are household members 16 years and older, the sample will be representative of the entire population because information about children is gathered in the household interview and children are sample members, however not interviewed personally. Further, children born or adopted by initial sample members will be included into the sample, and children reaching the age of 16 years will become respondents. This concept prevents the panel population from getting older on the average. Drop-outs by death are replaced by the following generation.

In addition, all sample members will be followed and reinterviewed within the area of the Federal Republic of Germany regardless of whether they move with the whole household or whether they split off. This concept allows observation of the process of splitting off and setting up of new households. Non-sample members who live together with sample members in later waves will be recruited as respondents if they are 16 years and older, but if these households split again, only the sample members, not the non-sample members, will be followed.

Special problems might arise for the sample of foreigners. At the moment, the government makes an effort to motivate the guestworkers to move back to their native countries. On the other hand, a lot of them still bring their families to Germany apparently expecting to stay. Which development will dominate is unknown. Movement out of Germany will be shown by the panel. To cover what is probably the most important part of the immigration, there is the rule that persons moving from foreign countries directly into a sample household will get the status of a sample member. This means that they will be followed, if they split from the sample household. Persons who come from foreign countries and form their own households or move into the institutionalized population are not represented in the panel; it may be necessary to include a sample of them from time to time in the panel. Altogether these rules will make it possible for the Socio -Economic Panel to reflect the natural development of the population, except for the small group of immigrants mentioned last.

Refusals are one of the most serious hazards for a panel. If there are too many, the panel will be destroyed in a short time. In the Federal Republic of Germany, there are no experiences with response rates which can be directly compared. In existing German panels, response rates have been far smaller than for instance for the Panel Study of Income Dynamics (11) or the National Longitudinal Surveys (12) in the U.S. But in most German studies, very little was done to increase the response rates. A comparison clearly shows that the more was invested into the motivation of the respondents and into the fieldwork, the better were the results. Therefore, for the Socio - Economic Panel special efforts are made to reduce refusals.

An intensive interviewer training is important for a successfull data collection. The procedure has already been explained above. To help the interviewers make the first contact, they presented a small booklet to the respondents which described the purpose of the study, how the household had been selected, why it is necessary to have its participation and how data protection works. Finally, it announced that all participants will receive a ticket for a well known lottery.

In general, the interviewers have to try to contact the households personally, making as many visits as necessary to reach somebody. If they get a refusal or can not secure a personal contact because the respondents do not open the door, a letter from the Institute including the booklet is sent to these households, unless they had refused for very serious reasons such as data protection. Then another interviewer tries to contact the household. If this attempt is unsuccessful too, a specially trained group of telephone interviewers will try to contact the household to get a date for an interview. About five months were calculated for the fieldwork of each wave, to allow enough time for the motivation of reluctant respondents and to wait for ill persons and persons on holiday.

After each data collection, the respondents will get a letter with a little present. After the first wave they got a ticket for the lottery. Some weeks before the next data collection starts, they also will receive a short report of the study with the announcement of the next wave. Beside the motivation, these two additional contacts function to up-date the address file (13).

Interim findings of the fieldwork for the first wave

The fieldwork for the first wave of the Socio-Economic Panel will be terminated in early October 1984. Therefore, at the moment, it is only possible to report some interim findings and to give an estimate concerning the expected response rates.

The fieldwork started with the interviewer training and the test interviews in February 1984, and with the main part of the data collection in March 1984. However, difficulties arose and led to a delay of the time table.

As mentioned above, a lot of interviewers did not qualify during the interviewer training. In several sample points this required to exchange interviewers and to start the training procedure again.

For sample B of foreigners, there were difficulties to receive for all selected sample points the addresses from the respective county administrations. In some cases, intensive discussions about data protection were necessary. For some sample points the addresses did not arrive in time, and, therefore, the fieldwork in these areas was delayed for several weeks.

The sample, as described above, was designed assuming an amount of unusable addresses of about 6 % and a response rate of about 70 %. The fieldwork showed that these expectations could not be realized, at least for the moment.

On the one hand, the response rates proved to be somewhat lower. Compared with other German surveys, there will be less drop-outs because of households that were not reached, which shows that the fieldwork was much more intensive than usually. But there will be a higher proportion of refusals. The first findings give the impression that the still virulent debates about data privacy had a negative effect. Many households refused with the arguments 'invasion of privacy' or 'this is another kind of conducting the census cancelled in 1983'.

On the other hand, especially for sample B, there were a lot of addresses unusable. Addresses registered at the administrations are sometimes absolete or wrong. With foreigners, this is a common problem: moved persons are registered at the new place of residence but not deleted at the administration of the previous location.

This situation required to add further addresses to the sample. The following procedure was used: For all sample points of sample A, where not enough interviews had been realized, the last address of the primary address selection was used as a new starting address and, again, every seventh household was selected. For sample B, 20 addresses had been drawn in each sample point, so that there were enough addresses left to fill up sample points.

The following table shows some interim findings about response rates. As still addresses are in the field, it is only possible to give an estimate of the expected results.

It is obvious that sample B will have much better response rates than sample A. This reflects previous experiences with data collection on foreigners. It is much more difficult to contact foreigners and to overcome their general suspiciousness. But if this point is passed, it seems to be easier to conduct an interview with them. A possible explanation is that foreigners are very afraid that the interviewer might be an official of the administration which possibly might send them back. If this problem is solved, foreigners are much more interested in talking about their life and less conscious of data protection problems. However, response rates of Italians are similar to the response rates of Germans, although the Italians are the best integrated group of the guest-workers. This is also a result of other German studies with foreigners.

Compared with the results of other panel studies, the interim findings for the first wave of the Socio-Economic Panel are satisfying. Obviously, the first waves are the most difficult to conduct. Although, the Panel Study of Income Dynamics has very high response rates in later waves, in the first wave in 1968, it was only 76 % (11), (12). The Swedish panel study on Household Market and Nonmarket Activities (14) could secure response rates of 85.2 % with the contact interview and 74.5 % with the personal visit, both in the beginning of 1984. But these studies are not really comparable, since in the Panel Study of Income Dynamics only the head of the household is interviewed, while in the Swedish panel mostly head and spouse are the respondents in the household, and only in some cases a third household member will be asked.

The Survey of Income and Program Participation (SIPP) has a very high response rate with 95.2 % in the first wave (15). The design of this study is similar to that of the German Socio-Economic Panel. All grown-up household members are respondents. But there is an important difference in the interviewing procedure. Because of the data privacy law of the Federal Republic of Germany, it was not possible, at least in the first wave of the German Socio-Economic Panel, to conduct proxy interviews. In consequence, only self-interviews were allowed. This was an extremly difficult situation, especially in large households. As the interviewers had to contact each household member until they got an interview or a refuse, in most cases several household contacts were necessary. Interpreting the response rate, this means that, depending on the subsample, for 61 % to 69 % of all households self-interviews with all household member 16 years and older have been conducted. In contrast to this, the SIPP allowed proxy interviews, and usually the interviewers did not have to visit the household several times to reach each respondent personally. For respondents who were not present at the time of the interview, a proxy interview with another household member was conducted. In consequence, the response rate of the SIPP of 95.2 % includes about 40 % proxy interviews and only about 60 % self-interviews. If one keeps in mind this difference in the interviewing procedure, the response rates of the German Socio - Economic Panel and the SIPP are much more similar.

Acknowledgements

The Socio - Economic Panel is conducted in team work by a group of scientists headed by Professor Dr. Hans-Jürgen Krupp. Findings are a result of joint efforts. I am grateful to the collegues of the German Institute for Economic Research of the Special Collaborative Programme 3 and of Infratest Sozialforschung for the good cooperation and the many fruitful discussions. I am particularly indepted to Professors Michael E. Borus, Heinz Peter Galler, James N. Morgan and Guy Orcutt for their many constructive comments on several early drafts of this paper, all of which led to improvements in the final version. Geil Orcutt was a most patient counterpart when discussing English terminology. Ingrid Wernicke did a wonderfull job when typing the manuscript. However, I am responsible for all remaining mistakes.

Footnotes

1) Saterdag, Stegmann 1980, 1982; Giesen et. al. 1981; Birk, et.al. 1978; Lewin, Schacher 1979, 1981, 1982; Projektgruppe Hochschulsozialisation 1979; Forschungsgruppe Hochschulsozialisation 1981, 1982; Dann et. al. 1978.
2) Pohl 1980, 1982; Herlth et.al. 1981; Rosenstiel et.al. 1983.
3) Büchtemann 1983.
4) In the German terminology, the word household is used for people living together and sharing the budget. Thus households can be composed of non-related persons.
5) The institutionalized population includes in the German definition inmates and staff living in an institution and not running a household of their own. Statistisches Bundesamt 1974, p. 13.
6) Statistisches Bundesamt 1974, p. 19. More recent data are not published.
7) In the German definition, families are couples or single parents, living together with their unmarried children (two-generations-families). Couples, separate living couples, widowed and divorced persons without children are also families. But single persons do not form a family. Statistisches Bundesamt 1982, p. 12.
8) The German Federal Statistical Office usually does not provide sample frames for non-administrative surveys as the US-Bureau of the Census does. Nevertheless, they tried to support the project.
9) Hanefeld 1982.
10) Duncan, Morgan 1981, p. 6.
11) A Panel Study of Income Dynamics 1982, p. 3.
12) Center for Human Resource Research 1983. The response rates for the different cohorts can be obtained from the Center, but the figures are only available beginning with the second wave.
13) It may be easier in Germany than in the U.S. to get new addresses of households that have moved, because there is an obligation to register the move with the local administration, although this system does not always work, especially with foreigners. In addition, very often the new address is known at the post office. If not, the interviewer will have the task of inquiring in the neighbourhood.
14) Klevmarken 1984.
15) Herriot, Kasprzyk 1984, pp. 17 - 18, figure 5.

References

A Panel Study of Income Dynamics 1982: Procedures and Tape Codes, 1981 Interviewing Year, Wave XIV, A Supplement, Ann Arbor.

Birk, Lothar, Heinz Griesbach, Karl Lewin, Martin Schacher 1978: Abiturienten zwischen Schule, Studium und Beruf - Wirklichkeit und Wünsche, HIS-Hochschulplanung Nr. 28, München - New York.

Büchtemann, Christoph F. 1983: Die Bewältigung von Arbeitslosigkeit im zeitlichen Verlauf. Repräsentative Längsschnittuntersuchung bei Arbeitslosen und Beschäftigten 1978-1982. Abschließender Gesamtbericht, in: Forschungsbericht Reihe Sozialforschung, Band 85, herausgegeben vom Bundesminister für Arbeit und Sozialordnung, Bonn.

Center for Human Resource Research (Hrsg.) 1983: The National Longitudinal Surveys (NLS), Handbook 1983 - 1984, The Ohio State University.

Dann, Hanns-Dietrich, Bernhard Cloetta, Gisela Müller-Fahrbrodt, Reinhard Helmreich 1978: Umweltbedingungen innovativer Kompetenz. Eine Längsschnittuntersuchung zur Sozialisation von Lehrern in Ausbildung und Beruf, Stuttgart.

Duncan, Greg J., James N. Morgan 1981: Persistence and Change in Economic Status and the Role of Changing Family Composition, in: Hill, Martha S., Daniel Hill, James N. Morgan: Five Thousand American Families - Patterns of Economic Progress, Vol. IX, pp. 1 - 44, Ann Arbor/Michigan.

Forschungsgruppe Hochschulsozialisation 1981: Anlage und Instrumente des Abiturienten-Follow Up 1976/77. Arbeitsunterlage 37 (Neuausgabe 1981), Konstanz.

Forschungsgruppe Hochschulsozialisation 1982: Wissenschaftlicher Bericht 1979 - 1982, Konstanz.

Giesen, Heinz et. al. 1981: Vom Schüler zum Studenten. Bildungslebensläufe im Längsschnitt, München.

Hanefeld, Ute 1982: Die 78er ADM-Stichproben. Eine kritische Beschreibung der bevölkerungsrepräsentativen Zufallsstichproben für die Bundesrepublik Deutschland, Sfb-Arbeitspapier Nr. 74, Frankfurt - Mannheim.

Herlth, Alois, Franz-Xaver Kaufmann, Joachim Quitmann, Regina Simm, Klaus Peter Strohmeier 1981: Generatives Verhalten oder Familienentwicklung? Werkstattpapier I, Institut für Bevölkerungsforschung und Sozialpolitik, Universität Bielefeld.

Herriot, Roger A., Daniel Kasprzyk 1984: The Survey of Income and Program Participation, SIPP Working Paper Series No. 8405, US Bureau of the Census, Washington D.C.

Klevmarken, Anders N. 1984: Household Market and Nonmarket Activities. The first year of a Swedish Panel Study. A paper presented at the annual meeting of the American Statistical Association, Philadelphia.

Lewin, Karl, Martin Schacher 1979: Studium oder Beruf? - Studienberechtigte 1976, 2 Jahre nach Erwerb der Hochschulreife. HIS-Hochschulplanung Nr. 33, München et. al.

Lewin, Karl, Martin Schacher 1981: Studienberechtigte 1978 - Studien- und Berufswahl im Wandel? HIS-Hochschulplanung Nr. 35, Hannover.

Lewin, Karl, Martin Schacher 1982: Studienberechtigte 1978 - Studienaufnahme, Studienfinanzierung, Zufriedenheit. HIS-Hochschulplanung Nr. 42, Hannover.

Pohl, Katharina 1980: Zielsetzung, Durchführung und Ergebnisse der Längsschnittuntersuchung des Bundesinstituts für Bevölkerungsforschung, in: Materialien zur Bevölkerungswissenschaft, Heft 16, Wiesbaden.

Pohl, Katharina 1982: Konzeption und derzeitiger Stand der Paneluntersuchung des Bundesinstituts für Bevölkerungsforschung zu Fragen des Familienbildungsprozesses, in: Zeitschrift für Bevölkerungswissenschaft, Heft 4, pp. 499 - 521.

Projektgruppe Hochschulsozialisation 1979: Anlage, Instrument und Durchführung der Studentenbefragung im WS 1977/78 (Studentenpanel), Universität Konstanz, Zentrum I Bildungsforschung/Sonderforschungsbereich 23, Arbeitsunterlage 56.

Rosenstiel, Lutz von, Günther Oppitz, Erika Spieß, Martin Stengel 1983: Motivation generativen Verhaltens - Ergebnisse. in: Lüer, G. (Hrsg.): Bericht über den 33. Kongreß der Deutschen Gesellschaft für Psychologie in Mainz 1982. Göttingen.

Saterdag, Hermann, Heinz Stegmann 1980: Jugendliche beim Übergang vom Bildungs- in das Beschäftigungssystem. Bericht über die Ergebnisse der Basiserhebungen einer Längsschnitt-Untersuchung, in: Beitr. AB 41.

Saterdag, Hermann, Heinz Stegmann 1982: Ausbildungs- und Berufsverläufe. Die IAB-Längsschnittuntersuchung in: Mertens, D. (Hrsg.): Konzepte der Arbeitsmarkt- und Berufsforschung. Eine Forschungsinventur des IAB, Beitr.AB 70, pp. 445 - 473.

Statistisches Bundesamt 1974: Fachserie A, Bevölkerung und Kultur, Volkszählung vom 27. Mai 1971, Heft 11, Bevölkerung in Anstalten, Stuttgart -Mainz.

Statistisches Bundesamt 1982: Fachserie 1, Bevölkerung und Erwerbstätigkeit, Reihe 3, Haushalte und Familien 1981, Stuttgart - Mainz.

The German Socio - Economic Panel

Interim findings concerning the response rates of the first wave [1)]

	Sample A	Sample B				
		Greek	Italian	Jugoslavian	Spanish	Turkish
Planned net sample size	4,500	200	300	300	200	400
<u>Interim findings:</u>						
Addresses initially used	6,847	312	461	467	353	629
Addresses added	1,139	57	179	88	46	67
Total addresses	7,986	369	640	555	399	696
Still in the field	278	30	30	32	26	41
Addresses already worked on	7,708	339	610	523	373	655
Not in the population or wrong addresses	430	65	136	119	105	102
Net addresses	7,278	274	474	404	268	553
Accepted interviews	4,425	184	299	284	188	368
Nonresponse	2,853	90	175	120	80	185
Expected sample size [2)]	4,545	201	308	300	200	393
Expected response rates [2)]	61 %	66 %	62 %	69 %	68 %	66 %

1) Interim figures from late September 1984. Final figures may change slightly since a small proportion of addresses is still in the field.

2) A substantial proportion of the addresses still in the field seem to be no valid household addresses. In many of these cases, the interviewer as well as the telephone contact was unable to reach a person. A final controll via the post office being conducted presently shows, that the person does not live there anymore, that it is a business address, etc. Therefore, the total expected number of accepted interviews will not increase according to the interim figures, but the number of net addresses will be somewhat reduced. However, the response rates will remain more or less unchanged.

HOUSEHOLD MARKET AND NONMARKET ACTIVITIES. THE FIRST YEAR OF A SWEDISH PANEL STUDY.

N. Anders Klevmarken, University of Göteborg, Sweden.

1. Purpose and scope.

The goal of the project Household Market and Nonmarket Activities (HUS) is to conduct basic research on the economic behavior of Swedish households. Economic research has long been concerned with how households use their resources, such as the days per year spent in gainful employment, the time spent in housework and leisure, the proportions of income which are respectively saved and spent and which goods and services the households use. All of this has direct importance for our welfare and for the economic development of the countries. However, in order to obtain additional results with this research we need better data about households. For instance, compared to its share of GNP the resources used to produce data about the household sector are small.

There is a large need to know how households adjust themselves to changes in their economic situation both when it improves and when it declines. Some examples of such changes are new taxation and welfare regulations, changes in municipal daycare, new forms of savings and new economic conditions for various kinds of housing. Cross-sectional data do not have much to offer for such an analysis, we do need longitudinal data.

Distributional issues have usually been analysed with cross-sectional data, but again, for the most interesting issues as, for instance, the duration of poverty, longitudinal data are needed.

Compared to previous research we want to attain above all three improvements with our project. The first is to study the entire household. Economic decisions are frequently a family matter. Individual decisions are influenced by the family situation and by other household members, and there are joint decisions about children, work, housing and leisure. The living standard and well-being of an individual depend very much on the household to which this person belong. For these reasons, we think it is essential to have data about the entire household to study people's economic activities and living standards in a realistic way.

Secondly, we intend to collect information which gives a complete picture of the household member's economic situation and planning. Then we can study how households' and their members' various activities such as gainful employment, housing, recreation, consumption and savings behavior depend upon each other and interact, and how they are affected by national and municipal government policies. An example is how the supply of day-care can affect employment and leisure activities for both spouses in the family. We thus need data which cover all these aspects of a household's economic activities.

Thirdly. we want to have longitudinal household data which will improve our understanding of households adjustments to economic policy, changes in their economic behavior and the duration of economic states.

The HUS-project will give new contributions in the following respects:

- it gives a total picture of the household's employment, purchases, savings, household and leisure activities,
- it will be the first investigation in Sweden in which time-use by households can be studied from the economic perspective,
- it gives significantly improved capability in studying how schooling, work experience, wages, family situation and childcare affect the level of employment in the household,
- it provides new potential in studying role division in couples with regard to employment, housework and raising children,
- it will provide previously unknown information on how much time households spend on the maintenance of their housing, cars, boats, etc.,
- it makes it possible to study which factors are important in the choice of type of housing,
- it makes it possible to study how economic welfare in the form of consumption standards and wealth are distributed among households.

A more detailed description of our problems is found in the original research program, Eliasson & Klevmarken (1981).

The HUS-project is a joint operation of researchers from the University of Gothenburg, The Industrial Institute for Economic and Social Research (IUI), Stockholm and the Stockholm School of Economics.

2. Population, general design and data collected.

The first wave of the HUS-study aims at the population of households residing in Sweden by the end of January 1984. Individuals living in institutions who did not form their own household and did not prepare their own meals were excluded. Because it is expensive to interview old people and the response rate would have become low households with very old members were also excluded.

The units of analysis are both the household and the individual household member. The definition of a household resembles that normally used in consumer expenditure surveys. Those who live in the same dwelling and regularly have meals together belong to the same household. Family members who temporarily live somewhere else are also included. Since each household will become interviewed three times during 1984 we will experience household splits and also new members joining and old household. The rules adopted for these situations were to follow all individuals selected for an interview in the first contact with the household unless they moved abroad or into an institution. New household members are not interviewed.

Since there is no sampling frame of households or dwellings but well a register of all residents of Sweden, each household was identified through a randomly selected individual. The household to which this individual belonged was included in our sample. It was, however, not feasible to interview all household members. Instead we decided on a scheme where the head of the household and his spouse were always interviewed.

If the randomly selected person was neither of these two, this third person was interviewed in addition. In this way we could ascertain some information about other adults in the household and also get a "clean" random sample of designated persons. Individuals below 18 years of age and above 74 i.e. born before 1910 or after 1966 were excluded from the sampling frame. The randomly selected persons were thus all in the age bracket 18-74, however other household members need not be. Children were not interviewed.

Each randomly selected person received an introductory letter, which was followed by a telephone contact. In this first contact the interviewer informed about the survey, asked for the name and age of each adult household member and their family relations. Finally, the interviewer also booked a time for a personal interview with each respondent. On the basis of the information obtained in this short interview one household member was designated household head. In a household with both spouses living together the husband was called head. Much of the information collected concern economic facts about the household which the husband on the average is expected to know more about than his wife. In households with two or more adult household members but without a married or cohabiting couple the person with the highest income was the designated head. Who was called head was never communicated to the respondents, but only used to decide who would get questions about housing and other issues not particular to any single household member.

The second step in the field work was a personal interview with each respondent, i.e. a maximum of three per household. This interview was planned for an average interviewing time of 60 minutes for the head and 45 minutes for other adults. The questionnaire included the following sections:

1. Family composition
2. Social background
3. Schooling
4. Marital status
5. Childcare
6. Health status
7. Labor market experience
8. Employment
9. Job search of unemployed
10. Not in the labor force
11. Housing and housing costs
12. Tenants
13. Real estate ownship
14. Cars
15. Boats
16. Other durable consumer goods
17. Incomes and assets.

The information collected in this personal interview included economic details about the household like housing expenditures, mortages and interest payments, and various income items and assets which most people would not be able to give without consulting note, bills and taxforms. Some respondents might also hesitate to reveal these data because of their sensitivity. For these reasons the respondents were asked to give this information in wrighting on a questionnaire which they put into an envelope and sealed before it was handed over to the interviewer. The interviewers were instructed to interview other household members while waiting for the questionnaire. In this way no sensitive information was revealed to the interviewer.

We had originally planned to obtain most of the information about incomes, transfer payments and personal wealth from government data registers via Statistics Sweden. The government gave us a permit to copy data from the tax assessment form of each respondent. However, to get access to these files we needed the social security number of each respondent. The Datainspection Board also required that we obtain the consent of each respondent to use register data. In a pretest it soon became clear that it was very difficult to get the social security numbers. In Sweden the most commonly used person-id in public and private data files is the social security number, and the public debate about computers and invation of personal privacy made respondents very reluctant to reveal their social security number. To investigate this further some 500 randomly selected respondents (not included in our HUS-sample) were asked if they preferred to give their social security number or would rather give the information we asked for directly in a questionnaire. About 27 per cent answered that they were willing to reveal their social security number while 65 per cent preferred to give it directly, 5 per cent spontaneously refused to do either while 3 per cent did not know. To avoid a high nonresponse we thus had to change our plans and it was decided that each respondent would have a choice, either to answer questions about income, assets etc. directly in writing, without revealing it to the interviewer, or to give the social security number and not answer the income questions etc.

A disadvantage with this scheme is that the data obtained from the questionnaires might not be fully comparable to the taxfile data. To minimize this problem direct reference was made in the questionnaire to the items of the tax assessment. Most of the field work was also done immediately after the tax assessment forms were submitted to the authorities. In our judgement this problem of comparability is minor compared to the nonresponse problem.

The possibility to compare the income and wealth estimates from register and interview data with the corresponding population totals can be used for nonresponse compensation and model validation, which to some extent makes the nonresponse problem less severe.

In addition to the personal interview, the respondents were contacted twice for two telephone interviews about their time-use and consumption expenditures. The method used was an adaptation of the yesterday question technique previously used at the ISR, the University of Michigan. It is perhaps best described as an one day retrospective interviewer administered diary. The basic idea is that the interviewer goes through the past 24 hours with the respondent and asks him or her to recall for each activity, when it started and ended and if the respondent made any expenditures at the same time. In addition to their time-use and expenditures the respondents were in each interview asked about their labor market status. In the last interview we also asked about purchases of durables since the beginning of the year, and how frequently they had used certain public services.

For each household two days were randomly selected from the 365 days following the 15th of February 1984. Ideally the time-use estimates should cover a calendar year to match income and wealth data, but all Swedes are busy with their tax assessment forms in the end of January and beginning of February and for this reason we wished to minimize our field work during this period and decided to start collecting time-use data when the tax assessment forms had been submitted by February 15.

A pilot study showed that designated dates for interviews caused a relatively high nonresponse because the respondents were not always available for an interview on the selected days. Thus, for each designated date there were also two alternative dates which could be used if a contact could not be reached on the first day. These alternative dates were selected on the same week day[1)] one and two weeks respectively after the designated date.

The interviewers were told to contact the respondents on the day following the designated day. If they could not get an interview they should try the next day and the next day again. If they still were unsuccessful they should repeat the same scheme for the first and second alternative day. They were not allowed to conduct interviews with a longer memory time span than for three days.

The personal interview should in general preceed the two telephone interviews. This plan was chosen because it is easier to explain a survey in person than by telephone and in this way the interviewer would not be a complete stranger to the respondents in the time-use interviews. For practical reasons we were, however, not always able to follow this plan. Some telephone interviews had to be made from the telephone unit of the survey institute SIFO in Stockholm. A few were also made before the personal interviews.

3. Pretest experiences.

In April and May of 1982 we made a rather extensive pilot study based on a random sample of 315 households from Western Sweden. There were five main purposes of this study, namely, to

a) compare different methods of collect-expenditure and time-use data,
b) get estimates of response rates and an idea of what might be important for the response,
c) test the questionnaires,
d) develop coding and editing procedures,
e) train the project staff in the entire survey operation.

The results from this pilot study have been reported in Klevmarken (1982, 1983). Here follows only a very brief summary.

There were altogether three contacts with each household. The first one was a short contact interview by telephone with a randomly selected person to establish the household composition and to ask a few demographic questions. Then two interviews followed with each respondent in each household. The same rules for designated respondents were used as explained above. One interview was personal and one was made by telephone. In addition, leave behind expenditure diaries were administered to each respondent and leave behind time-use diaries to a few respondents.

The response rate in the major contacts was as low as 50-55 per cent, which is much lower than we would find acceptable in a main study. In short, we ascribe this result at least partly to the ambitions design, the short timespan during which the field work had to be done and the budget constraints, which did not permit paying the respondents nor permitted expensive nonresponse follow ups. Our conclusion was that improvements in the design and use of response stimulating measures should make it possible to increase the response rate.

Our nonresponse analysis gave the following additional results:

- o The initial nonresponse was rather high. This was probably the combined effect of the following features:
  - (i) The survey was introduced by telephone rather than in a personal visit.
  - (ii) In this telephone interview we asked for family composition and previous marriages and living arrangements which some respondents might have found invasive.
  - (iii) When the interviewer concluded the interview by explaining the design of the study many respondents found the work load too high. This shows that the first interview should be in person and the telephone contact preceeding it should not be used to ask questions, only to make arrangements for the first interview.

- o A major drop in the response rate also occurred immediately after the contact interview, i.e. many respondents refused to keep an expenditure diary. Leave behind diaries tend to increase nonresponse. In this case a better result might have been obtained if the relative simplicity of the diary had been demonstrated by the interviewer in a personal visit. In the pilot study the diary was explained in the initial telephone contact and then mailed to the respondents.

- o Old respondents showed a relatively high nonresponse in those parts of the survey which involved relatively more work, i.e. diaries and long interviews about time-use. For this reason we decided not to include very old persons in the main survey.

- o Nonresponse was relatively high in urban areas.

- o There was no indication of a strong relationship between nonresponse and income or socioeconomic group.

- o Refusals made up a very large share of the nonresponse. This indicated that we would have to do a much better job in explaining the importance of the survey and also provide some personal stimulus to obtain a better cooperation.

Results from tests of alternative data collection methods can be summarized in the following way.

For almost all commodities the yesterday question technique gave smaller estimates of average expenditures than leave behind diaries. Since we have no reason to expect that leave behind diaries would give overestimates this result shows that yesterday questions in the form used in the pilot study tend to underestimate househould expenditures. However, in the main study we have improved the methodology by deleting a few supplementary questions for each activity, by adding after the time-use sequence a few questions about expenditures previously not mentioned and by giving stricter rules for how the questions should be asked. We thus hope that the underreporting problem is reduced.

Even if the yesterday questions will not give systematic errors, expenditures recorded only for a few days for each respondent give unreliable estimates. If the shopping pattern during the week is approximately the same for all commodities then it might be possible to adjust the sampling design to this pattern and in this way increase the efficiency somewhat. It is, however, not likely that this gain in efficiency would become so high that a longer observation period for less frequent purchases would not be needed. In the main study we have thus supplemented the last time-use interview with questions about purchases of major durables.

The yesterday question technique to collect time-use information has worked relatively well once the interviewers got used to it. The time-use questionnaire requires much more training than a traditional interview briefing gives. The pilot study did not include a comparison with the closest alternative, a self administered leave behind diary, but a comparative evaluation of these two methods would be useful for future data collection. All we have been able to do so far is to compare estimates of time-use in aggregate activities for the United States and Finland with our own estimates. There is a striking similarity in the time-use pattern between the three countries (Flood, 1983). We have also compared the response to yesterday questions with that of retrospective questions covering two weeks. Similar to results from other studies we found that retrospective questions for a longer period tend to give systematic errors. Time-use for less frequent activities was underreported compared to the results from yesterday questions.

The pilot study also included a comparison between estimates of time off work at work from the time-use diary and from direct questions about time normally used for meals and coffee breaks, personal errands etc. Longer hours was on the average reported for "normal" time off work at work than for the corresponding activities from the yesterday question diary.

If telephone interviews could be used instead of personal interviews was another issue investigated in the pilot study. Our experiences show that a difficult and demanding study like ours should be introduced to the respondents in person. If not, the nonresponse rate is likely to increase. For respondents we could, however, find no significant difference in time-use or expenditures between interviews made in person and those made by telephone.

4. Sample design.

The sample was obtained by a two-stage cluster design. The clusters were stratified by a rather unconventional procedure. First, Sweden was divided into nine major geographical areas. Within each area a cluster analysis was run on the zipcode areas to group them into strata. For each zipcode, variables were used measuring the age, income and occupational distributions, the share of the population living in owner occupied houses and the share of foreigners. The zipcodes were stratified in four to eight strata depending on geographical area.

The primary selection unit was, however, not the zipcode but commune or more precisely, those zipcodes which belong to a commune within a stratum. We preferred to use this primary selection unit rather than zipcode to reduce the geographical dispersion of the sample. If we had used zipcode we would have needed more interviewers than we could possibly recruite and train. From each stratum two primary units were selected.

In the second sampling stage a number of individuals were chosen from each primary unit of zipcode areas. As a sampling frame we used the SPAR register, which is a register of all residents of Sweden. We were aiming at a self-weighted sample with about 10 individuals in each primary unit. With a sample size of 2200 this required a little less than 200 interviewers, which approximately is the size of the interviewer staff of the SIFO institute which did our fieldwork.

In practice it was not possible to get a completely self-weighted sample. The population figures of the SPAR register were about half a year old and the effective sample size of each primary selection unit deviated a little from 10. In all we obtained 2131 individuals. We thus used a sample of 2131 households and expected to interview on the average 1.65 respondents per household, totaling approximately 3500 respondents.

Our budget did not permit a larger sample although it would have been desirable to increase its size. Our possibilities to analyze subgroups of households will now become somewhat limited. For instance we will get relatively few unemployed, academicly trained or wealthy people.

To make feasible an inference to the annual time-use of the Swedish population we decided to use a random day design. This design should ideally take advantage of seasonal and weekly variations in time-use. Efficiency calculations made for the time-use studies at ISR, the University of Michigan (Karlton [1983]) indicate that a sample of two days might be sufficient. The marginal increase in efficiency of additional days decrease rapidly. If it is desirable to draw conclusions about a particular season or, for instance, calculate within individual variance estimates for weekdays, more than two days are needed. Budget considerations, however, limited our alternatives considerably. We could not afford more than a sample of two days for each respondent. Given this constraint, differences in time-use between seasons and between workdays and holidays were built into the design in the following way.

The calendar year was split into a winter and a summer season and the days of a week were divided into workdays and weekends or holidays. The 365 days of a year were thus grouped into four strata.

	Workday	Weekend & holidays
Winter season	A	B
Summer season	B	A

The sample of households were randomly divided into two halves A and B. For each household in group A a winterseason workday and a summer season Saturday, Sunday or holiday were randomly drawn and for each household in group B a winter-season Saturday, Sunday or holiday and a summer season workday. In each stratum days were drawn with equal probabilities and without replacement.

We also tried to balance the sample of days with respect to the strata of clusters and to the two selected primary units from each stratum.

5. Fieldwork considerations.

In addition to the normal interviewer training which SIFO provides its interviewers each interviewer had to participate in an one day training session particular to the HUS-project. In all, about 20 sessions were hold around the country. During these sessions the interviewers were informed about the purpose and scope of the project and about the details of the field work procedure. They also practiced with the HUS-questionnaires. Before the field work started each interviewer also had to do testinterviews in the field with the questionnaires for the head of the household. These questionnaires were sent in to SIFO, reviewed and corrected and returned to the interviewers. Those interviewers who were not able to produce satisfactory testinterviews did not participate further.

From our pretests we know that nonresponse is our major problem. Since the present study is to become the first wave of a future panel study it is essential to have a high response rate. For this reason it was decided to give the respondents some kind of renumeration. Our budget did not allow for cash payments but we have tried a combination of gifts and a lottery. At the and of the personal interview the interviewer gave the respondent a list of a few small gifts, for instance, one item was a chocolate box. The respondent could either chose one of the gifts from this list or participate in a lottery to win a flight ticket for two persons to Paris. This design does not only serve the purpose to reduce nonresponse, but it will also give interesting data about peoples choice behaviour under uncertainty. Each respondent will also receive a summary of the main results from the survey.

6. Response.

When this is written the fieldwork with the contact interviews is completed and only three personal interviews remain, while most of the telephone interviews still have to be done. Table 1 shows the field work logg by the 19th of June which exhibits the nonresponse by reason as of this date.

No interviews were attempted with those labled "Not in the population". The response rate for attempted contact interviews is 75.5 per cent and for attempted personal interviews 74.6 per cent. Nonresponse is almost entirely due to refusals. One might also note that there was almost no increase in nonresponse in the telephone interviews following the personal visit.

7. Concluding remarks.

After the first year of fieldwork we will have data about labour market status from three different interviews for each responding household member. This will become our first piece of longitudinal information. Our plans are to return to the same households in the beginning of 1986 to obtain data on changes in household composition, and new data on housing, labour force participation, earnings, incomes and assets. New time-use and expenditure data will probably not be collected until later.

Footnotes.

1) If the designated day was a workday (holiday) but any of the corresponding weekdays one and two weeks later was a holiday (workday) the closest following workday (holiday) was chosen as an alternative day.

References.

Eliasson, G. & Klevmarken, N.A. (1981), Household Market and Nonmarket Activities. Research Program and Proposal. IUI Research Report No 12, 1981.

Karlton, G. (1983), Sampling Design Issues in Time Diary Studies, Appendix B to A Proposal to The National Science Foundation for support of A 1984-85 Study of Time Allocation Among American Households. Principal Investigator: F. Thomas Juster. Co-Principal Investigators: Martha S. Hill and Frank P. Stafford. The University of Michigan, March, 1983.

Klevmarken, N.A. (1982), Household Market and Nonmarket Activities (HUS). A Pilot Study. Research Report 1982:3, Department of Statistics, University of Gothenburg.

Klevmarken, N.A. (1983), Collecting Data for Micro Analysis. Experiences from the HUS-pilot Study. Paper presented at the Symposium on Microanalytic Simulation Models to Support Social and Financial Policy, in Bonn, June 20-22, 1983.

Table 1. Non-response by reason August 20, 1984.

	Contact interview	Personal visit	Telephone interview I	Telephone interview II
Not in the population				
Living abroad	21	27		
Moved, address unknown	35	41		
Not Swedish speaking	25	55		
Unhealthy, institution	16	22		
" at home	23	37		
Other	20	42		
Subtotal	140	224	230	221
Nonresponse				
Refusal, personal integrity	29	54		
" too busy	31	67		
" never participates	17	33		
" participated before	19	37		
" interview workload too high	12	25		
" not interested	35	57		
" after attempted personal persuasion	165	285		
" other	123	275		
Not available	57	63		
Other	0	0		
Subtotal (of which additional non-response in the telephone interviews)	488	896	886 (29)	867 (7)
Accepted interviews	1504	2635	1758	396
Still in the field	0	0	881	2271
Total sample size	2132	3755	3755	3755
Response rate (%)	75.5	74.6		

Note 1. The first column shows the number of randomly selected persons which equals the number of households. The last three columns show the number of individuals designated to participate. For those households which did not complete the contact interview the number of designated respondents is unknown. These households have in columns 2-4 been enumerated as single person households.

Note 2. The response rate in the last row of the table is the ratio between the number of accepted interviews and the total sample size less those not in the population.

THE AUSTRALIAN NATIONAL LONGITUDINAL SURVEY

Ian McRae, Bureau of Labour Market Research

1. Introduction

The Australian National Longitudinal Survey (ANLS) is essentially a survey of the labour market, and the data it collects will focus principally on explaining the workings of the Australian labour market. The survey is based on a large probability sample. It will collect data which relates a variety of areas not usually covered in a single collection. This, in conjunction with its longitudinal nature, will make it a very valuable source of data for Australian researchers and policy advisers.

When the survey is in full operation it will include 12 000 people selected from two sampling frames. Of these 3 000 will be selected from a list of registered unemployed and 9 000 from an area based sample. Data will be collected on a wide variety of issues, but with a main focus on the labour market and unemployment. The scope of the survey is restricted to people aged 15-24 in 1984, in order to allow a detailed analysis of this group.

This paper has been structured as follows. Section 2 briefly discusses the background to the decision to conduct a longitudinal survey. Section 3 describes how decisions regarding broad design issues were taken, based on the principal objectives of the survey. Section 4 describes some of the secondary objectives, which include the use of the data in evaluation of Government labour market programs. Evaluation is discussed in more detail in Section 5. Section 6 outlines the data required to meet these objectives. Sections 7, 8 and 9 describe the sample design and estimation procedures proposed for the survey. Section 10 notes briefly some practical aspects of data collection, and Section 11 makes brief comment on the methods of analysis likely to be used.

2. Background

In January 1983 the Australian Government was considering a range of means of assisting the long term unemployed. In the course of the deliberations they became aware, not only of the dearth of information relating to this group, but that certain key issues could not be resolved from the traditional cross-sectional data available. Subsequently funds were allocated for the conduct of a longitudinal survey of at least 3 years with the primary purpose of learning more about the long-term unemployed, and the Bureau of Labour Market Research (BLMR) was asked to conduct the survey.

It should be pointed out that Australia is generally very well supplied with labour market data. The government statistical agency, the Australian Bureau of Statistics (ABS) conducts a monthly labour force survey of a national sample of 32 000 households. This survey encompasses all the conventional labour market data, and for 8 or 9 months of the year includes a supplementary survey. The supplementary surveys cover a wide variety of labour market and social issues, including alternative working arrangements, child care, labour mobility, labour force experience, income and working conditions.

In addition to the monthly survey, ABS conducts a series of Special Supplementary Surveys. These surveys cover, for example, household expenditure, housing, health and handicapped persons. They tend to be much larger and more complex than the supplements to the monthly survey and they may be used to focus on particular sub-groups within the population.

The monthly labour force survey is not well suited to longitudinal analysis. Month to month gross flows data are published. However the mobility of the population is such that attrition is quite high (10% of the population cannot be matched from month to month), and the rotation pattern is such that no one is in sample for more than 8 months.

There have been a number of longitudinal surveys conducted in Australia, but none on a large scale, in the sense of covering large numbers of both people and data items. Most of the more prominent studies have been based on school leavers, with an initial interview at school, and mail follow ups at different periods (see Blandy and Richardson 1982, Williams 1981, and Dowling and O'Brien 1979). These surveys have been extremely valuable, but it has not been possible for them to closely control attrition nor to ask a wide range of questions. There have been no substantial longitudinal studies of the unemployed or the wider community.

March 1983 saw a change in Government in Australia, with the new Labor Government being committed to large scale job-creation schemes. In developing these schemes, however, the policy makers had requirements for more information than was available about the unemployed. This led to confirmation of interest in longitudinal data and the initiation of a major longitudinal survey directed at data on policy issues.

The first wave of the survey (which encompasses only one quarter of the total sample) will go into the field in September 1984, with the introduction of the remainder of the sample being planned for 1985. The current plans encompass surveys in 1984, 1985, 1986 and 1987.

3. Broad Design Issues

Given the origins of this survey, the initial objective was to consider the following questions:

- who are the long-term unemployed (not only in demographic terms but in terms of many other aspects of their background)?
- how do they become long-term unemployed (or conversely, which of those who become unemployed get jobs most quickly)?
- how do people change with increasing durations of unemployment - with respect to job search methods, physical and mental health, finances etc.?

There are a number of ways in which these objectives can be met, and a number of different international 'models' were considered. Two key issues were involved in these considerations - firstly the general form of the survey, and secondly (given the general form chosen) the choice of the cohort to be surveyed. These are addressed in turn below.

The first approach considered was to select a sample from the current stock of unemployed, and follow their experiences through time (see for example White 1983). However this approach has three deficiencies. First, it does not allow us to examine the process of becoming long-term unemployed, although this information is vital in determining policies to influence this process.

Second, this approach does not provide a control group. While population information from the Census and Labour Force surveys provides suitable comparative information for some issues, for others these sources provide inadequate detail. Third, after a number of years the sample will no longer consist solely of unemployed people. The value of this eventual sample structure is far from clear. A survey was therefore required which provided adequate samples of long and short-term unemployed, employed and not in labour force.

There are a number of major surveys which have this coverage. The most prominent of these are the US National Longitudinal Surveys (NLS) (see, Borus 1981) and the Panel Study of Income Dynamics (PSID) (see, Morgan 1974). The former uses national samples of individuals within particular age/gender cohorts. The latter is based on a sample of families, and includes people of all ages and both genders.

The principal objectives of this survey relate to unemployment and other issues of relevance to the labour market, and hence to the behaviour and experience of individuals. It was therefore considered that an approach which was based on individuals was more appropriate. This does not mean that information on families will be neglected, and in fact a large amount of data about other family members is being collected. It does mean however that the individual is the sampling unit, and the unit which is to be tracked.

The next issue in the broad design was to consider whether the survey should cover all age/gender groups, or be constrained to a single cohort. In Australia about half the unemployed and about half the long-term unemployed are young (under 25). Any attempt to sample uniformly across all ages would therefore mean half of the sample would be spread across a forty year age range, allowing little scope for detailed analysis of any particular subgroup.

After considerable debate on both technical and policy grounds it was decided that the survey should concentrate on a single youth cohort, defined by the age range 15-24 years. As noted earlier this group comprises about half the unemployed, and until recently it accounted for nearly 90 per cent of government expenditure on training and employment programs for the unemployed. Examination of this group has the added advantage that it allows examination of the process of transition from school to work, and covers the age range within which a majority of Australians marry and many have their first child.

The sample will cover both people who are long-term unemployed and those who are not. This can most easily be done by selecting an area sample of the defined age group. Only about 10 per cent of such a sample would be unemployed, however, and about half of these long-term. This is inadequate for analytic purposes, so there was a need to over-represent the unemployed. This could be done by oversampling areas likely to contain unemployed persons, as is done in the US NLS (CHRR 1982). This would increase the numbers of unemployed in the sample not only in the first wave but also in later waves, assuming areas retain their employment characteristics.

An alternative was to select a list sample of known unemployed in the first survey. Although this may not retain such high unemployment levels in the sample in later years as the first approach, the number of unemployed in the first wave would be controlled, and the number of people moving into long-term unemployment in the second wave is maximised. There are two extensive lists of unemployed people in Australia, that held by the Commonwealth Employment Service (CES) and that held by the Department of Social Security (DSS). For reasons of both convenience of access and greater comprehensiveness, the CES list is to be used. The use of this list also provides a sample which represents the registered unemployed and hence provides an opportunity for analysis of the activity of the employment service. Finally, the use of a relatively large list sample provides an opportunity to experiment with 'tracking down' individuals. As there have been very few substantial longitudinal surveys in Australia, there is very limited experience in this process.

The outcome of these considerations was a sample of:

3 000 persons aged 15-24 who had been registered with the Commonwealth Employment Service (CES) for at least 3 months. (Group 1)

9 000 persons selected from an area sample which covers all but very sparsely settled areas of Australia. (Group 2)

(Sample sizes are discussed in Section 7.)

The survey is to be implemented as follows:

		Group 1	Group 2
September 1984	:	First Wave	
June-Aug 1985	:		First Wave
September 1985	:	Second Wave	
June-Aug 1986	:		Second Wave
September 1986	:	Third Wave	
June-Aug 1987	:		Third Wave
September 1987	:	Fourth Wave	

The separation of the first waves of the two parts of the survey was made for various reasons, which were all essentially practical

- relatively little time and staff resources were available for survey development
- the survey was to be conducted by a commercial agency (the reasons for this are discussed below) and even the small survey was a large job by their standards
- it was considered necessary to test the quality of this work, particularly with respect to tracking, before launching the major part of the survey
- as the principal focus of the survey was to be unemployment a relatively large and detailed study of the unemployed was of considerable value in its own right.
- there was considerable policy interest in obtaining data as early as possible, so it was desirable to begin the survey in 1984.

4. Survey Objectives

As noted in the previous section, the prime focus of analysis from this survey will be long-term unemployment. The processes of entry and exit to this state and the changes observable in people as unemployment duration lengthens will be examined. However there are many other potential uses for the data. The organisations with whom there has been substantial contact include the Department of Education and Youth Affairs, whose main interest is in those who leave education 'early'. The Commonwealth Employment Service, is also interested in using this survey to examine the use of, and attitudes to, its services.

The survey will allow a great deal of analysis of labour market, educational and social welfare issues. These include:

- analysis of flows between employed/ unemployed/not in labour force states; to answer the question of whether there are many people unemployed for a short time or a few people unemployed for a long time
- examination of the process of transition from school to work
- examination of patterns of usage of the Commonwealth Employment Service
- study of usage of government employment and training programs and their long-term outcomes
- studies of links between poverty and unemployment - who are the poor? do they stay poor?

Given the very high cost of the survey, which could be as high as one and a half million dollars per year in its peak years, its potential needs to be exploited as far as possible. Thus while the survey cannot be thought of as an omnibus to be used freely, and data will only be included which can benefit from the longitudinal nature of the data base, it is intended to provide access to it as a research vehicle for both government and non-government research.

5. Government Program Evaluation

A particular objective of the survey is to provide information for evaluation of government employment and training programs. There are two aspects to this use. For most programs, participants will be asked a very limited set of questions about their participation in programs and the benefits they perceive from this participation. The largest job-creation program in Australia, the Community Employment Program (or CEP), will be treated differently. The evaluation of this program was initially envisaged as drawing on a separate longitudinal survey, as a principal evaluation issue was whether participation enhanced long-term employment outcomes. However, there is substantial commonality in data requirements, and the timing proposed for the evaluation and the longitudinal survey (ANLS) were similar. In the interests of efficiency (particularly with respect to survey development and management) the two surveys have been merged. As far as interviewers are concerned, there is one survey, with different question streams for CEP participants, and the rest. As far as analysts are concerned there are two surveys - the ANLS and the CEP evaluation. The ANLS will serve as a control group for the evaluators in some contexts. The full structure of the first wave of the survey is shown in Table 1 attached.

The discussion which follows relates

to the ANLS only, but the actual survey development and management have been substantially influenced by the need to provide an appropriate framework for CEP evaluation.

6. Data Requirements

The data requirements for this survey are broad. The initial auspices of the survey was as a basis for policy oriented labour market research, but there is considerable interest from many sections of government and academia. The data are constrained by a requirement that the interview be restricted to about one hour. This restriction is somewhat arbitrary, but was based on experience of surveys in this country. This suggested that interviews which run longer than an hour are more likely to experience severe 'respondent fatigue', which may lead to eventual refusal to complete the interview.

The central data requirements of the survey were defined by the objectives outlined in Sections 3 and 4. These were refined and added to in an exhaustive round of consultations with potential users of the survey data in many arms of government and in the wider research community. The list of data items in Table 2 attached reflects these consultations.

Thus the basic demographic data for the respondent and their family will be collected, plus detailed educational information including an attempt to assess "on-the-job" training. Detailed employment histories are to be collected for the period covered by the survey, and the 12 months preceding the first interview. Broad data on the three years preceding this period will also be collected. Current labour force data will be collected consistent with the concepts definitions and patterns used by the Australian Bureau of Statistics in their Special Supplementary Surveys, plus data from the unemployed on the methods they use to seek work and the wage they seek. Limited data on health, both physical and mental, will be sought, plus quite detailed income data, and somewhat limited assets data. Several questions will also be asked about satisfaction with working conditions.

Participants in Government employment and training programs will be asked about these programs, with a special more detailed set of questions for CEP participants. The unemployed will be asked about their usage of, and satisfaction with, the Commonwealth Employment Service.

It is proposed to limit the questions on attitudes, aspirations, and expectations. It is proposed to ask about attitudes to women working and about educational plans, though it is not proposed to ask about occupational expectations.

7. Sample Size

Sample size considerations invariably involve compromise between 'needs' and cost/management considerations. The 'need' in this case is to be able to examine small sub-groups. For example of 3 000 sample about:

- 1 400 will be unemployed 9 months or more
- 500 of these will be female
- 250 of these will be aged 15-19
- 125 will be in their first year out of school

Examining this particular group of young females with respect to their educational levels, whether they live with parents, income, employment experience, country of origin etc is difficult because of the sample sizes. This is, however, clearly a group with quite specific needs and will be studied.

With the 9 000 area sample there will be reasonably large numbers still at school initially, who can be 'tracked' as they enter the workforce. However, between year one and year two of the area sample it is expected that about

- 80 will leave school aged 15
- 220 will leave school aged 16
- 250 will leave school aged 17

The 15 and 16 years olds will not have completed secondary school, although some of the 17 year olds will have completed. The analysis of the eventual labour market outcomes of some of these groups by gender and birthplace will be difficult because of the very small sample sizes.

Thus, sample sizes were not established in a highly scientific way, but were rather seen as a trade-off between a survey size which could be managed on one hand, and 'adequate' samples of important sub-groups on the other. While expected standard errors of important variables were calculated, considerations such as those shown above played the major role in sample size considerations. The examples indicate that the final sizes chosen are not excessive.

8. Sample Design

The sample design for this survey is constrained by the usage of the Commonwealth Employment Service (CES) register as the framework for the first stage. This register is not a unitary file, but rather a series of files held in each of 250 CES offices. The only practical sample design is therefore a multi-stage design, selecting first the CES offices, and second the persons registered at those offices.

The sample is to be allocated proportionately to the six States and two Territories, as the main requirement is for national rather than state level data. Within States the sample was stratified into metropolitan/non-metropolitan areas, with further stratification of the large metropolitan areas. The smaller metropolitan areas and non-metropolitan areas were carefully ordered, and not divided into strata. These steps led to 20 strata.

The sample will be selected in two and, when relevant, three stages. Firstly CES offices will be chosen and then people registered at the offices. If this group is too geographically dispersed (as can happen in country areas in particular), a sub-selection of a geographical cluster will be made. The sample will consist of 60 clusters each of 50 persons.

The area sample will be based on the same geographic regions, both for reasons of minimising interviewer travel and to provide a suitable control group for the sample of registrants. While the level of clustering is yet to be finalised, and it may prove practical to spread the sample further, the current proposal is to select 10 blocks with an average of 15 respondents from each area controlled by a selected CES office.

The selection of the area sample will involve a screening phase, but as it is expected that, on average, one 15-24 year old person will be found in every second house this will not be excessive. All eligible persons found in the screening are to be interviewed.

As a consequence there will be households where more than one person is included in the sample. This approach was chosen as it will give samples of both husband and wife and also of siblings. It is also cheaper, but will have a higher variance for some estimates, and response problems may arise if too much time is requested of the household. This last problem has not yet been tested in the field, and is the one aspect of this plan which may require reconsideration.

The sample, once selected, is to be followed over a number of years. There is very little Australian experience of tracking in household surveys and it is not clear what levels of sample attrition are likely. American experience of sample attrition is very varied. The NLS older cohorts have retained between 68 per cent and 78 per cent of their sample after 10 years (see O'Neill, 1983). The new youth cohort has maintained 96 per cent of the original sample for each year of its conduct so far (Rhoton, 1984). The University of Michigan Panel Study of Income Dynamics lost 14 per cent of the original sample between the first and second waves, but suffered very small losses thereafter (see Survey Research Centre, 1983). In the survey of the unemployed conducted in the UK, the retention from 1980 to 1981 was 78 per cent (White 1983).

Australia has quite high levels of mobility among the young, particularly the unemployed, but does not have the same range of public registers as countries like the USA. The main tracking mechanism will be the names and addresses of family and friends supplied by the person at the initial interview. Beyond this there are phone books and electoral rolls, neither of which are likely to be up to date - particularly with respect to the group of interest. There are some possibilities of using other government registers, but these vary from area to area and have not yet been fully explored. Every effort will be made to retain contact with respondents in the period between surveys but until the process is actually begun it is not known how successful this will be.

It is intended to make every effort to retain contact with respondents in the period between surveys but until the process is actually begun it is not known how successful this will be.

It is intended to make every effort to minimise attrition. People anywhere in Australia will be included in the sample, although in more remote areas this may be by telephone. People who leave Australia will be omitted from the sample. Each year efforts will be made to contact the full original sample, so respondents may miss out one year and be re-included the following year.

As noted above, no other Australian study has had the resources to follow up respondents in the way which is proposed for this survey. What the final attrition rates will be is strictly a matter for conjecture at this stage.

9. Estimation Procedures

Estimation procedures will be considered further when the pilot tests have been conducted. The notes below indicate current thinking.

Basically, estimation will always relate to the population at June 1984. The simplest version will be for the 1984 survey, which will be a list sample only. Within each stratum the numbers of persons registered by age, sex and duration of registration will be available. Selection probabilities (say Phi) for each person (i) selected in stratum h will be known. Response rates within certain groups will also be known.

The most important groupings expected to show differential non response are those who do and those who do not move house between selection and interview. Information on marital status and country of origin will also be available. In the discussion which follows it is assumed that moving is the main cause of non-response. If this is not the case the same methods would apply.

Assuming that all sampled persons can be identified as 'movers' or 'non-movers', the sample will be post-stratified into these categories.

Separate estimation and non-response adjustment would then be undertaken for each post-stratum. Information on age, sex, and duration of unemployment will be available for the complete population. As both response rates and attributes are likely to be related to these factors, this data may be used for benchmarks.

The current plan, which will be finalised using the results of the pilot test, is that within each selection stratum (a) separate estimates (using conventional unequal probability estimators) would be made for 'movers' and 'non-movers'. (b) These estimates would then be adjusted for non-response.(c) Further adjustments would then be made to known age-sex-duration of unemployment benchmarks.

When the area sample is introduced there will be several variations to this approach needed. Armstrong (1978) provides a very thorough discussion of methods for estimating from multi-frame surveys. The following discussion briefly describes the methods to be applied in this case. Firstly, probabilities of selection will change. In particular, all people who were eligible for selection in the list sample will have two opportunities to be selected - in the list sample and in the area sample. Thus there will be three groups of people in the final sample

i) Those selected in the list sample.

ii) Those selected in the area sample but eligible for the list sample.

iii) Those selected in the area sample and not eligible for the list sample.

While the probabilities for group iii) may be calculated directly, those for groups i) and ii) are more complex. In both these cases the total probability of selection, (where h denotes selection stratum) is:

Prob (selected) = Prob (in list sample) + Prob (in area sample) - Prob (in both samples)

$= P_h + P_h^A - P_h^{**}$

Given that the area sample in any region will be three times as big as the list sample, and denoting by U the unemployment/ population ratio in the area, and f the proportion of unemployed people in that region included in the list sample, it can be shown that:

Prob (person i is selected) = $(1 + 3u(1-f))P_h$

Thus at one extreme, with an unemployment/population ratio of 15 per cent and f of .025

$P(i \text{ selected in } R) = 1.45\ P_h$

More likely numbers are 10 per cent and .06, giving

$P(i \text{ selected in } R) = 1.28\ P_h$

In some rural areas f will be very large, perhaps as large as 0.5. In this case, if u = 0.1

$P(i \text{ selected in } R) = 1.15\ P_h$

Thus the joint probabilities can be quite variable depending on the situation.

The actual estimation procedure for the area sample will obviously be different to the list sample in the first year, as the notion of 'movers' will not apply. Straight forward age-sex adjustment at the smallest geographic level for which benchmark data is available will be used. This will probably be at the state level, possibly classified into metropolitan and non-metropolitan. In later years, separate non-response adjustment for movers and non-movers would also be applied.

As noted earlier, this approach indicates current thinking. If further investigation reveals response rates are much the same for movers and non-movers the process could be simplified. The number of age-sex-duration cells is yet to be considered, but may be very few. Adjustments which are proposed for stratum level could be undertaken at state or even national levels to reduce complexity. These issues have not yet been thoroughly addressed.

Given the clustering of the sample and the complexity of the estimation process, variance estimation will clearly need to be of a replicated nature. Exactly how it will be done is yet to be decided, although balanced repeated replication techniques (see McCarthy, 1966 or Kish & Frankel, 1970) are suitable.

10. Fieldwork and Data Processing

Data collection and processing of the ANLS will be undertaken by a commercial market research company. As yet there are no large scale agencies in Australia like the social survey research organisations in America. The larger Australian market research companies do however undertake social research, and have substantial experience of such work.

Efforts are being made to involve the Federal Government's official statistical agency, the Australian Bureau of Statistics (ABS) in later stages of the survey. The ABS is the largest and most competent survey organisation in the country, and while it has not been able to participate so far, may be in a position to provide field work and other assistance in the future.

The survey is planned as purely personal interview. There may be some use of the telephone for people who have moved to particularly remote places, but this would be quite restricted. There is also some possibility of switching to telephone interviewing in later years, but the survey until 1987 has been planned and costed on the basis of personal interview.

The consultant will be contracted to provide BLMR with a clean data tape, and data processing will be undertaken at

BLMR. It is proposed to manage the data files using the SIR data management system, and to undertake most of the analysis using SAS. It is hoped that even variance estimation will be possible in SAS, as programs for Balanced Repeated Replication have already been published.

11. Analysis Proposed

In Sections 3 and 4 a series of issues are noted which BLMR and other agencies will wish to address in a variety of ways. Table 3 attached lists a variety of other issues for research which will be given priority. As it is planned to make the data base publicly available as early as possible it is likely much analysis will be progressively undertaken outside BLMR.

The initial publications will be aimed at providing a description of the population being sampled, and those aspects of their history to be collected by recall. As longitudinal data becomes available even relatively simple descriptive statistics will contain a great deal of information. In the initial analysis it is not proposed to use higly sophisticated techniques. These techniques will no doubt be considered at a later stages and in the contexts of particular research projects.

12. Conclusion

The Australian National Longitudinal Survey is an ambitious and expensive undertaking. It will provide a wealth of data which has not been available previously in Australia, both longitudinal and cross-sectional, and has the capacity to make a substantial contribution to both policy making and research in this country.

References

Armstrong, B. (1979), Test of Multiple Frame Sampling Techniques for Agricultural Surveys, Survey Methodology, Vol. 5, No. 2, Statistics, Canada.

Blandy, R and S. Richardson (19820, The Fate of the Class of 71, The Australian Bulletin of Labour, Supplement No. 4.

Borus, M.E. (ed) (1981), Pathways to the Future, Center for Human Resource Research, The Ohio State University, Columbus, Ohio.

Centre for Human Resource Research (1982), The National Longitudinal Surveys Handbook, College of Administrative Science, The Ohio State University, Columbus, Ohio.

Dowling, P. and G. O'Brien (1979), A Longitudinal Study of the Work Values and Job Expectations of School Leavers, Flinders University of South Australia Working Paper No. 36.

Kisch, L and M.R. Frankel (1970), Balanced Repeated Replications for Standard Errors, Journal of the American Statistical Association, Vol 65, No33.

McCarthy, P.J. (1966), Replication: An Approach to the Analysis of Data from Complex Surveys, National Center for Health Statistics, Series 2, No. 14, Washington D.C.

Morgan, J.N. et al. (1974), Five Thousand American Families, Institute for Social Research, Ann Arbour, Mich.

O'Neill, J. (1982), Review of the National Longitudinal Surveys, Department of Labor, Washington D.C.

Rhoton, P. (1984), Attrition and the Longitudinal Surveys of Labour Force Behaviour: Avoidance, Control and Correction, IASSIST Quarterly Vol. 8, No. 2.

Survey Research Centre (1983), User Guide for the Panel Study of Income Dynamics, University of Michigan, Ann Arbour.

White, M. (1983), Long-term Unemployment and Labour Markets, Policy Studies Institute, London.

Williams, T. et. al., (1981), School and Work in Prospect, ACER Research Monograph No. 101, Melbourne.

Table 1. Survey Structure 1984

Sample Component	Age 15-24 years	25+ years
ANLS	Random sample of those registered 3 months with CES (3 000)	
CEP - Participants	Random sample of CEP participants - constrained by geographic clustering of ANLS (500)	Random sample of CEP participants - constrained by geographic clustering of ANLS (500)
CEP Control	Matched sample of registrants who are eligible for CEP positions (500)	Matched sample of registrants who are eligible for CEP positions (500)

Total sample size for 1984 - 5 000 persons

Table 2. Data Items: Broad Headings

- Demographic variables relating to the respondent and other residents of the respondent's household
- Family Background
- Marital History and Fertility
- Education
- Impediments to Employment
- Current Labour Force Status - Job search methods and reservation wage
- Job History - last 12 months
- Work Experience Prior to '12 months' before
- Health - Physical and Psychological.
- Participation in and Satisfaction with Government Programs
- Income and Assets
- Satisfaction with Working Conditions

Table 3 Selected Issues for Research

- To determine the concomitants of the lengthening of unemployment spells. Are all socioeconomic, age, and gender groups equally at risk?
- To identify the labour market consequences of long-term unemployment. Are youth permanently disadvantaged after long-term unemployment or does this experience have a positive impact on later labour market outcomes?
- To determine the changes occurring in an individual's attitudes and behaviour over the duration of a spell of unemployment. Issues considered include finances, housing, family formation, educational expectations, retraining and occupational and geographic mobility.
- To investigate the differences in incidence and impact of long-term unemployment and frequent short spells of unemployment. Information is needed on the social and demographic differences in those experiencing the two types of unemployment, on the effects of both types on later labour market success, and on the effect of the business cycle on these patterns of unemployment.
- To examine whether the incidence of long-term unemployment is different for males than it is for females. To examine attributes such as education, work experience and geographic mobility in relation to male and female labour force experiences.
- To investigate different school to work transition patterns, the relative importance of factors involved in successful transition.
- To examine the role of vocational training in labour market experience, especially as regards trades careers.
- To gain insights into the ways in which permanent occupations are found. There are conflicting hypotheses about the relatively frequent job changes by youth. Evidence is needed to assess the causes and effects of these job changes.
- To assess the impact of different job search methods on the duration of job search, job acceptance, and subsequent earning levels and job satisfaction.

☆U.S. GOVERNMENT PRINTING OFFICE: 1985-461-105/20019