

John F. Coder, Bureau of the Census

INTRODUCTION

The Income Survey Development Program (ISDP) was the research and development phase of the planned Survey of Income and Program Participation (SIPP). The ISDP was intended to examine and resolve operational and technical issues for SIPP and provide a strong framework for the ongoing survey for the 1980's. Work began in ISDP as far back as the spring of 1977. Since then, three major survey operations have been conducted. In addition, a special survey of recipients of several government transfer programs and high income households selected from IRS records will be completed in August 1980.

Indeed, considerable money and time have been devoted to SIPP development. Ultimately SIPP is planned as a national longitudinal household survey. Mainly under the sponsorship of the Department of Health and Human Services, the survey will collect data on cash and in-kind income, program participation, eligibility, personal history, wealth, and other related topics. These are indeed difficult subjects on which to collect reliable information. Not only is the data collection phase extremely difficult, the processing of longitudinal data and creation of a longitudinal data file are enormously complex tasks as we have discovered in the ISDP studies.

The purpose of this paper is to discuss and document some of the important aspects of the final ISDP effort, the 1979 research panel. The scope of the paper covers a selected cross-section of information about our experience with the survey. It concentrates on problems relating to noninterviews, nonresponse, and a presentation of some estimates of demographic and economic characteristics of the population derived from Interview Number 1. As of the time of writing of this paper, all interviewing of the 1979 panel will have nearly been completed. Unfortunately, due to the complexity of the questionnaire and data collection procedures only results from the first interview period and partial results from other periods can be included.

Results from the 1979 ISDP are heartening in some respects and somewhat disappointing in others. While nonresponse and noninterview rates were overall very acceptable given the nature of the survey, estimates of income recipients for some of the key transfer programs appear low. Some estimates of demographic characteristics also differ substantially from other sources such as the March 1979 CPS. Large standard errors on the ISDP make interpretation of results difficult in many instances.

SAMPLE DESIGN

The sample design and selection for the 1979 ISDP was certainly a key aspect of the survey operation. To improve the efficiency of the sample a scheme of oversampling of low and high income households was employed. This procedure was to allow for more reliable estimates for the areas of the income distribution of most interest to analysts. Oversampling of low and high income

households was achieved by using the Survey of Income and Education (SIE) as the major portion of the sampling frame. The income data collected for SIE sample households was used to stratify the sample frame into 9 strata ranging between "Under \$2,500" to "\$74,000 or more." In these two strata virtually all SIE cases within ISDP PSU's were taken. The sampling rate for the other strata varied declining toward the middle income brackets. In addition to the SIE, portions of the sample were selected from the Bureau's current surveys program reserves and from new construction permits. The above 3 sources yielded the general population portion of the sample of about 9,300 households. A second portion of the sample, the administrative frames, was selected from administrative sources and yielded sample cases for known SSI and Basic Education Opportunity Grant (BEOG) programs recipients. About 1,000 SSI disabled and 1,000 BEOG cases were selected to provide a nationally representative sample within the 130 ISDP PSU's.

DATA COLLECTION

The data collection for all of the ISDP operations have been carried out by the Bureau of the Census. Interviewing for the 1979 panel began in February 1979. A rotating interview scheme was used whereby one-third of the total panel was interviewed each month. The reference period for income and labor force questions was the 3-month period prior to the month of interview. Because of the longitudinal nature of the survey, households were visited at 3-month intervals throughout the 17-month period between February 1979 and June 1980. Two-thirds of the sample were visited six times and the other third five times.

A very detailed effort was made throughout the survey to develop and maintain an accurate profile of income reciprocity. In all, 50 different sources of income could be individually identified. Monthly amounts received for each type were recorded. In addition, 14 different asset types and income flows from them were recorded. The sale of these assets, where appropriate, was also enumerated.

In the initial interview an experiment was conducted involving three different collection procedures. There were two versions of the questionnaire. One version used household response to obtain income reciprocity while the other used individual questions on a person-by-person basis. The household response technique was tested in hopes of reducing interview time. In this version, a single question such as "Did anyone in this household receive Social Security?" was used followed by "Who was that?" in order to identify specific individuals.

One-third of the sample was interviewed using the household response technique. The other two-thirds of the sample were divided in half. One-half was interviewed using the person-by-person approach to reciprocity and required self-response for each individual. The remaining third of the sample was interviewed using the person-by-person approach but allowed proxy responses.

EVALUATIONS AND RESULTS

The 1979 ISDP was and continues to be a true learning experience. There are so many issues surrounding the sampling, data collection, processing, and evaluation of the ISDP for 1979 that we could not begin to cover them all. Included in this paper are a selected group of issues which are very basic to SIPP, to the collection of income data in household surveys, and to the design and testing of questionnaires in general.

COMPLETE HOUSEHOLD NONINTERVIEWS

Of major concern to ISDP and the future of SIPP was the ability to gain and maintain voluntary cooperation from sample households for a survey which was advertised as dealing with a subject as sensitive as income. The noninterview rate is usually a good indicator of success or failure in this area. The noninterview rate represents the percent of all occupied housing units for which no interview could be obtained, either because of a refusal or inability to make contact with the residents. The overall noninterview rate for the first interview of the 1979 ISDP was 7.5 percent. The refusal rate was 4.4 percent. This noninterview rate was somewhat higher than the comparable incoming rotations of the Current Population Survey (CPS) during the same period. The CPS had a noninterview rate of 5.8 percent and a refusal rate of 1.9 percent.

The noninterview rate for the first interview in a longitudinal survey is an important indicator; however, equally if not more important is the final noninterview rate. This final rate is a key because it indicates how willing respondents are to endure the long rigorous interviews covering many sensitive areas. After four interviews the 1979 ISDP noninterview rate had climbed to 12.4 percent. Still two interviews remain for two-thirds of the sample and one interview for the other third. This 12.4-percent rate compares to about a 10-percent rate for the 1972-73 Consumer Expenditure Survey which consisted of five quarterly interviews over a 15-month period.

Based on this experience for 1979, it seems that we can achieve reasonably low noninterview rates in a longitudinal income survey. Obviously the 12.4-percent rate is high and methods for reducing the attrition of households need to be found and implemented; however, given the oversampling of high and low income households from SIE and the nature of the survey, the noninterview rate is probably lower than most analysts believed possible.

PERSON NONINTERVIEWS

The ISDP allows for completely noninterviewed persons within a household for which some persons were interviewed. For the first interview period of February, March, and April about 1 percent of the persons interviewed in the area frame portion of the sample were noninterviewed persons. Over 53 percent occurred in the first month of interviewing and over 56 percent on the self-response panel.

Some demographic information was collected for the noninterviewed persons within interviewed households. The distribution of noninterviewed

persons by demographic characteristics appears to differ significantly from interviewed persons. As indicated in Table 2, males were more likely to be noninterviewed as were children. Reflecting the higher proportion of children as noninterviews is the higher proportion for persons 16 to 20 years old (probably students away from home at school). Older individuals are less represented in the noninterview universe of persons because many live alone and a noninterview for them is a complete household noninterview.

The high proportion of noninterviews for the self-response only panel results from rules which prohibited taking proxy responses. In some cases, proxy interviews were obtained for reluctant respondents or respondents who are temporarily absent such as students away at school. Many of these persons were probably incorrectly classified as person noninterviews when proxy interviews should have been accepted. The high proportion of noninterviewed persons in the first month of interviewing probably indicates some first month jitters or misunderstanding concerning procedures on the part of new interviewing staff.

HOUSEHOLD VS. PERSON-BY-PERSON RESPONSE TO INCOME RECIPIENCY

One of the more interesting aspects of the 1979 survey was an experiment involving two very different methods of collecting a profile of the income sources received for each individual 16 years old and over in the household. The household method used for one-third of the sample asked a single question to identify receipt of a specific income type or group of related income types by any household member. Affirmative responses were then followed by a question to identify exactly which household member received that type of income. For example, one question was "Did you or any member of this household receive Social Security Income from the U.S. Government?" If the respondent said "Yes," a followup question was asked, "Who was that?" The second approach to collecting income reciprocity profiles was the person-by-person approach. For this method each individual (or proxy) was asked about each income type. In either case the amounts of income were recorded after the receipt of the income type had been identified.

Our preliminary analysis without the benefit of standard errors does not indicate any startling differences between the two approaches as far as estimates of income recipients by type is concerned. Anticipated large standard errors will make a meaningful, statistical evaluation difficult or impossible.

INTERVIEW TIME

Another of the major concerns with the ISDP/SIPP survey is respondent burden. This burden is generally measured in terms of interview time although the true burden may also be a function of the respondent's perception of the time, the subject matter involved, and the accuracy required. In the ISDP/SIPP this burden is extremely important because it requires repeated cooperation of respondents over a long period of time. Failing to maintain respondent cooperation over the entire survey period because the burden in terms

of interview time is too great would require reduction in interview times. Although we do not know how interview time is related to respondent attrition in ISDP, we do have some valuable information concerning the interview times for the first interview period.

Interview times are shown in Tables 3 and 4. The data in Table 3 are for all households in the first interview period. Overall, median and mean interview times were 42 and 45 minutes per household, respectively. Some of the interviews extended to over two hours, but only about 2 percent. About 78 percent were under one hour. One person households averaged 28 minutes per interview whereas households having four or more adults averaged about an hour and 20 minutes each.

The data shown in Table 4 show mean interview time by interview money, by questionnaire type, and by the collection procedure for income flows from assets. Interview time declined significantly over the first 3-month interview period. In the first month, February, the average time was about 53 minutes per household. This declined in the March and April interviews to 44 minutes and 42 minutes, respectively. This reduction in average interview time probably reflects increased interviewer experience with the questionnaire and with the survey in general. There was no major differences in interview times by treatment for collecting income flows from assets. On average, households in the half sample for which flows were collected took 46 minutes. This compares to about 44 minutes for other households for which no income flows from assets were collected. Interview time by questionnaire type shows that the person-by-person approach took, on average, 47 minutes for both the self-response only panel and the regular response mode. The household response panel required an average of 42 minutes per interview to complete, a savings of about 5 minutes overall.

ITEM NONRESPONSE

Beyond total household noninterviews and person noninterviews we have a third source of missing data and an important aspect of nonsampling error, that is, item nonresponse. Item nonresponse has been a particularly troublesome problem with respect to the collection of income data in household surveys. Most of our experience with income nonresponse rates for the CPS income supplement is one of the reasons that the Census Bureau was interested in the ISDP/SIPP program. Based on our experiences with the CPS many analysts felt that treating the income data as less than the highest priority subject matter was a major factor contributing to nonresponse.

Nonresponse rates for selected income items on the 1979 ISDP interview period No. 1 questionnaire are shown in Table 5. These rates are based on item nonresponse for interviewed persons and do not include noninterviews or noninterviewed persons.

We have not had the opportunity to study these data in Table 5 in detail. In general, however, there appears to be several issues of importance or which will certainly require additional study. First, the difference between nonresponse rates for self and proxy respondents for hourly wage amount seems very large. It appears that data

quality for this crucial item is very dependent on the respondent. Second, there is a very high nonresponse rate for the salary or draw which self-employed persons pay themselves. The 52 percent nonresponse rate for this item is a very serious problem and needs to be studied.

ESTIMATES OF INCOME RECIPIENTS BY TYPE

One of the key objectives of the ISDP effort was to design a survey which reduced both the undercounting of income recipients and the misclassification of income. These have been two major problems with the March CPS supplement. Some preliminary results of the 1979 ISDP are shown in Tables 6 and 7. Without the aid of standard errors it is difficult to interpret these results, however, analytically, there are several important findings.

1) Federal SSI.--The performance of this survey for identifying recipients of Federal SSI seems to be much better than the March CPS. It is not known exactly what caused the improvement. It is known that the March CPS has significant misreporting of Federal SSI as Social Security or State SSI. It may be that these problems have been solved to a large extent by the ISDP approach.

2) Aid to Families with Dependent Children (AFDC).--We have a real problem with AFDC. Compared to administrative data we accounted for only about 62 percent of the recipients for January 1979. We know of several factors which are affecting this estimate. The first is an apparent severe undercount of single-parent families headed by women on the 1979 ISDP. The ISDP estimate of 7,515,000 compares to a March CPS estimate of 8,458,000. Since almost all AFDC recipients are women heading single-parent families, the underestimate of this group could have a significant impact on the AFDC recipient estimate. The second factor is probably related to misclassification problems. In Table 7 we have indication of an actual overcount of General Assistance recipients, a fact which indicates that perhaps some AFDC recipients did not know the correct source of their public assistance and therefore classified it as General Assistance rather than AFDC. In addition to problems associated with General Assistance there may be other misclassification problems as well. There were a total of 316,000 cases reporting "other cash welfare" or "unknown cash welfare." It is certainly likely that many of these cases were, in fact, AFDC, however, no investigation has been made to determine which of these cases are most likely to be AFDC. Finally, there is always the possibility that there is some hidden error in the data processing system which is causing the problem. We have been and will continue to search for such a problem.

3) Unemployment Compensation.--Unemployment compensation has been one of the more seriously underreported income types on March CPS. In general, the March CPS obtains about 60 percent of the total annual dollar amount paid according to administrative sources. We are encouraged by the count of January 1979 recipients which indicates that the survey estimate was about 85 percent of a comparable estimate derived from administrative sources.

4) Interest Income.--The income source showing the largest impact compared to the March CPS was interest income recipients. The ISDP methods for identifying recipients yielded an estimate of 98.4 million interest income recipients. This compares to a figure of 67.7 million from the March 1979 CPS. The ISDP estimate is restricted to persons receiving interest from savings accounts while the CPS number covers all sources of interest income for the calendar year.

5) Veterans' Payments.--Our long standing problem with significant underestimates of Veterans' payment recipients in CPS appears alleviated in the 1979 ISDP. The ISDP estimate stands about 88 percent of the administrative estimates of disability compensation and pension recipients. A common problem in the March CPS has been the misclassification of Veterans' payments as military retirement.

ESTIMATES OF AGGREGATE CASH BENEFITS

Survey estimates of aggregate cash benefits from government transfer programs are a function of both the number of recipients and the amounts reported by the recipients identified. Table 7 also contains estimates of aggregate cash benefits for selected transfer programs for January 1979. The data indicate some significant underreporting of aggregates for many of the income types. There are also several instances of overestimation. Both Federal SSI and Social Security estimates appear acceptable and show some improvement over March CPS annual figures when they compared to annual administrative aggregates. Overreporting of State SSI and General Assistance are related to misclassification problems involving Federal SSI and AFDC, respectively. The low estimate of AFDC recipients (63 percent)

appears to be accompanied by underreporting of amounts. The aggregate cash benefit is only about 51 percent of the administrative figure of \$891.5 million. The underestimate of unemployment compensation of about 17 percent is near the underestimate of 15 percent in recipients. Both Veterans' payments and Federal retirement have significant underestimates of the January aggregate and display problems similar to those encountered in the March CPS.

CONCLUSIONS

The only real conclusion we can make comfortably about the 1979 ISDP results is that much more analysis will be required before we understand where we are and how we got there in terms of the collection of income data in a quarterly longitudinal survey. We have obtained some extremely useful information concerning the expected magnitude of noninterview problems. These problems appear manageable given the 1979 interviewing scheme and special problems encountered for this particular panel. We have computed some preliminary item nonresponse rates and found some interesting differences when these rates are examined by self and proxy response. One of the major objectives of the survey was improvement in the reporting of income data. For some income types we have apparently made some improvements in comparison to the March CPS, for some, no real improvement is evident, and for still others we find some disappointing results. Interview time, although lengthy, appears to be nearly in line with the OMB guidelines of 20 minutes per person. Interview times were reduced significantly as the interviewers gained experience with the complex ISDP questionnaire.

TABLE 1.--HOUSEHOLD NONINTERVIEW RATES FOR THE 1979 ISDP BY INTERVIEW PERIOD

Interview Period	Noninterview Rate (Percent)	Refusal Rate (Percent)
1.....	7.5	4.4
2.....	10.3	7.2
3.....	12.4	9.4
4.....	13.6	11.5
5.....	13.8	11.7
6.....	12.4	11.1

TABLE 2.--INTERVIEWED AND NONINTERVIEWED PERSONS WITHIN INTERVIEWED HOUSEHOLDS BY SELECTED CHARACTERISTICS: 1979 ISDP INTERVIEW PERIOD NO. 1

Selected Characteristics	Interviewed Persons	Noninterviewed Persons
<u>Sex</u>		
Total.....	100.0	100.0
Male.....	48.1	59.2
Female.....	51.9	40.8
<u>Relationship to Reference Person</u>		
Total.....	100.0	100.0
Reference person with own relatives.....	41.8	29.3
Reference person with no own relatives.....	(A)	(A)
Spouse.....	34.3	18.1
Child.....	17.2	42.6
Other relative.....	3.5	3.7
Other nonrelatives.....	3.1	6.3
<u>Age</u>		
Total.....	100.0	100.0
16 to 20 years.....	12.7	26.1
21 to 24 years.....	9.4	13.7
25 to 34 years.....	21.1	17.7
35 to 44 years.....	15.3	8.5
45 to 54 years.....	14.2	19.3
55 to 64 years.....	12.8	9.4
65 years and over.....	14.3	5.3

(A) excluded from the distribution since the noninterviewed person category is not applicable to persons living alone.

TABLE 3.--INTERVIEW TIME BY NUMBER OF ADULTS 16 YEARS OLD AND OVER: 1979 ISDP INTERVIEW PERIOD NO. 1

(Numbers in thousands)

Time (minutes)	Total	Number of Persons 16+						
		1	2	3	4	5	6	7+
Under 15.....	806	806	-	-	-	-	-	-
15 to 30.....	17,644	11,051	6,223	322	34	14	-	-
30 to 44.....	26,015	6,456	16,910	2,250	372	27	-	-
45 to 59.....	16,632	1,724	10,572	3,104	982	207	43	-
60 to 74.....	8,334	577	4,602	1,751	1,170	207	27	-
75 to 89.....	4,968	68	1,839	1,627	876	346	170	42
90 to 104.....	1,824	70	488	452	496	219	33	66
105 to 119.....	882	24	228	299	482	70	76	3
120 to 134.....	626	-	245	46	160	167	8	-
135 to 149.....	278	13	20	11	72	101	11	50
150 to 164.....	156	-	-	99	17	3	14	23
165 to 179.....	93	-	14	19	29	-	31	-
180 or more.....	102	1	3	19	34	44	1	-
Not reported.....	1,931	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)
Median time.....	42	28	43	57	71	86	87	102
Mean time.....	45	28	44	60	74	90	97	116

(NA) Not available
 - Represents zero

TABLE 4.--MEAN INTERVIEW TIME PER HOUSEHOLD FOR SELECTED SAMPLE SUBGROUPS: 1979 ISDP INTERVIEW PERIOD NO. 1

Selected Subgroup	Mean Interview Time Per Household (Minutes)
<u>Month of Interview</u>	
February.....	52.5
March.....	44.3
April.....	42.2
<u>Questionnaire Type</u>	
Person, self response.....	47.3
Person, regular response.....	46.6
Household response.....	42.0
<u>Asset Flow Treatment</u>	
Income flows from assets collected.....	46.0
Income flows from assets not collected.....	44.4

TABLE 5.--ITEM NONRESPONSE

Item	(Percent)		
	Total	Self Response	Proxy Response
Hourly Wage Rate.....	10.4	5.1	23.5
Social Security Income.....	7.2	5.4	16.6
Federal SSI.....	6.6	4.2	14.6
Pension and Retirement.....	14.1	11.4	31.2
Nonfarm Self-Employment....	52.0	49.2	60.1
Net Rental Income.....	36.3	33.3	43.8

TABLE 6.—COMPARISON OF ESTIMATES OF INCOME RECIPIENTS BY SELECTED TYPES OF INCOME: THE MARCH 1979 CPS, THE 1979 ISDP 3-MONTH REFERENCE PERIOD, AND THE 1979 ISDP ESTIMATE FOR JANUARY 1979

(Numbers in thousands)

Selected Income Types	March 1979 CPS (1978 annual recipients)	1979 ISDP 3-Month Period	1979 ISDP January 1979 Monthly Estimate
Federal SSI.....	2,908	3,821	3,821
State SSI.....	615	322	220
Social Security.....	28,779	28,783	28,179
Aid to Families with Dependent Children... ^{1/} ...	3,247	2,238	2,150
Other public assistance... ^{1/} ...	1,178	1,151	1,162
Unemployment compensation...	5,409	4,236	2,905
Worker's compensation.....	2,128	1,267	862
Private pensions.....	5,843	6,181	6,074
Military retirement.....	1,252	1,194	1,194
Federal retirement.....	1,091	1,206	1,204
State or local retirement...	1,660	2,221	2,179
Interest income ^{2/}	67,752	98,418	(X)

(X) Not applicable.

^{1/} For ISDP other public assistance includes General Assistance, Emergency Assistance and Other or unknown cash welfare.

^{2/} For ISDP number is restricted to those with savings or credit union accounts. CPS includes all sources of interest.

TABLE 7.—PRELIMINARY COMPARISONS OF RECIPIENTS AND AGGREGATE INCOME ISDP VS. ADMINISTRATIVE RECORDS: JANUARY 1979

INCOME TYPE	RECIPIENTS (THOUSANDS)			AGGREGATE (MILLIONS)		
	ADMIN ESTIMATE (1)	ISDP ESTIMATE (2)	(1)-(2)	ADMIN ESTIMATE (3)	ISDP ESTIMATE (4)	(4)+(3)
FEDERAL SSI.....	3,833	3,805	.993	551.1	501.0	.902
STATE SSI.....	259	220	.849	15.6	20.9	1.340
SOCIAL SECURITY*.....	29,264	28,179	.963	8,148.2	7,627.5	.936
UNEMPLOYMENT.....	3,400	2,905	.854	1,011.5	842.5	.833
AFDC.....	3,451	2,150	.628	891.5	451.5	.506
GENERAL ASSISTANCE.....	630	791	1.255	100.9	115.5	1.145
WORKER'S COMPENSATION.....	(NA)	818	(NA)		336.2	
VETERAN'S PAYMENTS.....	4,600	4,013	.872	740.0	614.0	.830

*16 YEARS OLD AND OVER