THE GOALS AND OBJECTIVES OF THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

Charles A. Lininger, Department of Health and Human Services

The Survey of Income and Program Participation (SIPP) is a national longitudinal survey program of the Departments of Health and Human Services (HHS) and Commerce. The survey will collect information through household visits and link it with data from various administrative record systems to provide a comprehensive data base with unique analytical potential for government policy research, and for academic and business researchers. The first SIPP panel (a series of interviews repeated at intervals with the same persons and families) will begin in January, 1982.

Goals and Objectives

The major goals of the SIPP are:

(1) to extend the scope and precision of policy analyses for a wide range of federal and state tax and social welfare programs;

(2) to improve current estimates of income and income change, including annual and sub-annual estimates, by source of income; and

(3) to broadly assess the economic well-being of the population.

The fulfillment of these goals requires an extensive data base collected over time for the same families and individuals, including information on assets, money and in-kind income, other categorical eligibility criteria, labor force participation, and household composition. The longitudinal nature of the survey makes it possible to measure changes in these characteristics. To achieve these objectives, it is necessary that:

- o information be timely and readily available,
- o information be capable of being linked with administrative record systems such as the wage and benefit data of the Social Security Administration,
- o the program be flexible enough to supplement both the national sample with special samples and core questionnaire modules with topical questions, and
- o the program have an effective sample size with continuity of data collection which would provide reliable estimates of changes in the major variables over time, annually and sub-annually.

Uses of the Data

The Bureau of the Census: The income data from SIPP are expected to become a major source of the Census Bureau's income distribution and poverty statistics for general statistical purposes. Thus one purpose of the survey is to make accurate and reliable estimates of levels, size distributions, and changes in income.

To measure poverty, information is needed on the condition of families over time, distinguishing between temporary and longer-term poverty-level incomes. SIPP measures must therefore combine short-term measures (e.g., monthly, quarterly) with a capacity to aggregate to longer-term estimates (annually, biannually).

To accomplish these ends it will be necessary to make a comprehensive assessment of the economic status of the population which is multidimensional in scope. Along one dimension, it must combine objective factors of well-being (income, assets, program and labor force participation, household composition) with subjective factors (attitudes and expectations). Along another dimension, it must generate static, point-in-time estimates, and also longitudinal estimates and analyses of change over time based on repeated observations of the same individuals.

Department of Health and Human Services: Improved annual and subannual income data (especially total income and transfer income by type) also are needed by the Department of Health and Human Services, and especially by the Social Security Administration, to analyze benefit-related issues concerning federal income security and social welfare programs and selected state programs, as well as tax incidence and yield under varying assumptions. Some examples to show how SIPP data will be used to address policy issues and to measure the effect of policy changes are:

- o Determining how changes in eligibility rules or in benefit schedules affect both recipients and federal spending levels;
- o Assessing whether federal assistance is distributed equitably under present programs and procedures;
- o Determining to what extent benefits are received by persons and families who simultaneously participate in several government programs (e.g., Supplementary Security Income, Old Age, Survivors and Disability Insurance, Food Stamps, Medicaid, Unemployment Insurance), and whether the combined levels of benefits are adequate, or excessive;
- o Estimating how many persons and families currently eligible for benefits are not receiving them, and assessing the increases that would be expected in program costs if participation levels increased among such persons;
- o Determining how alternative taxing arrangements would affect income and payroll tax incidence and revenues;
- o Understanding the reasons for changes in benefit status, employment, and household membership, and recommending what should be done to moderate undesirable changes;
- o Recommending appropriate indicators of program target groups, exploring improvements in the measurement of economic hardship and poverty, and showing how the inclusion of selected assets and in-kind income would vary the number and types of families and persons in poverty;

- o Understanding how attitudes and knowledge of program alternatives affect program participation and other economic behavior, how income adequacy norms vary by family size and size of place of residence, and how they respond to secular trends in objective income.

Other Users: The information collected by the SIPP will be useful to analysts in other federal agencies and will be made available to them according to their interests. The variety of users and uses encountered to date are numerous.
Selected examples include:

- o analysis by the Department of Agriculture of asset holdings and Food Stamp recipiency.
- o analysis by the Department of Labor as a response to the recommendations in the Levitan Commission Report (1979) for new data and study of labor market-related hardship measures, and for study of changes in labor force and dependency status through time; and
- o analysis by the Treasury Department, as input to a modified version of the Personal Income Tax model of the Office of Tax Analysis.

Academic and business researchers have consistently expressed great interest in the SIPP program and are on record as desiring and expecting early access to the data. The SIPP sponsors agree to provide documentation and data access at the earliest possible time.

Major Features of SIPP

Four major aspects of the SIPP program are described separately in the following four sections-the content of household interviews, matched administrative records, design features, and reliability requirements.

1. Household Interview Content. The SIPP will link data from household interviews with information from administrative records (SSI, OASDI, etc.). The exact nature of the administrative data included in the SIPP data base will vary according to the program and policy issues being addressed. This section therefore emphasizes the content of the household interviews and concludes with a discussion of administrative data requirements.

Survey questionnaires will include a comprehensive set of questions on the social, economic, and demographic characteristics of the sample households, individuals, and various sub-household groups (e.g., families; program units) within them. This information will be contained within a data file that will support accurate estimates of characteristics of the population. Measures of changes in these characteristics and (where possible) causes or antecedents of these changes are of particular importance to assess levels of economic well-being and to provide data for policy analysis and simulation. While some flexibility in content is a part of the SIPP design, the main "core" data elements required include program participation, eligibility determinants, cash and noncash income, respondent attitudes and expectations about economic well-being, current labor force participation, household composition, and personal history.

1a. Program Participation:

A major objective of the survey is to provide better information about the participants in a wide variety of transfer programs providing cash and noncash benefits. To accomplish this, reliable information is needed on the receipt, value, and classification of benefits, a need that coincides with the general goal of SIPP to improve income reporting. Particular attention will be given to the measurement of multiple participation (e.g., joint receipt of Aid to Families with Dependent Children, Food Stamps, Unemployment Insurance, and/or Medicaid).
Combining these program participation data with detailed measures of other social, economic, and demographic characteristics (from both survey and administrative record data) will provide the necessary base for policy simulation and program analysis.

1b. Eligibility: One of the most difficult problems in analyzing and simulating program activity is to estimate and describe the number of persons, families, or households who are eligible for benefits from a variety of programs. The SIPP intends to collect the detailed information necessary to simulate eligibility. This includes data on categorical determinants (e.g., age, disability status, and family composition) and means-tested determinants (e.g., income, assets, and selected expenditures). Because eligibility requirements differ substantially from program to program and, for some, by state, the survey cannot collect all possible information needed to determine eligibility precisely for every program. SIPP will provide information from surveys and administrative records to make reliable and accurate estimates for the major government programs.

1c. Money Income: The survey is designed to measure and classify the income received by each adult in each sample household. Similar but less detailed information will be obtained for children under 16 during interviews with a responsible adult. The goals of the survey include improvements in the following areas:

- o Accuracy of reporting--SIPP questionnaires and interviewer training are intended to improve significantly the quality of income data over that which is now available from the March supplement to the Current Population Survey. Special attention will be directed toward resolving problems of misreporting, underreporting, and nonreporting of cash transfers and property income (e.g., interest, dividends, and rent).
- o Classification of income--detailed information will be collected on specific sources of income and on reasons for receiving such income. In the developmental stage of SIPP, progress was made in preparing detailed transfer payment profiles by improving the classification of income types that are frequently confused by respondents (e.g., Supplemental Security Income and Social Security Income). Improvements in classification will be pursued further in the SIPP program.

- Time-periods—data will be collected on sub-annual reference periods, in most cases monthly, although certain income types will be difficult to obtain for such time periods because respondents don't often receive or calculate them on that basis. For example, profits or losses for farm or non-farm businesses probably can be obtained quarterly, at best, and interest income, net rental income, and dividend income is difficult to obtain on a monthly basis. Income data will be accumulated into annual estimates for each individual.
- An "annual round-up" interview, following the traditional tax-filing deadlines, is planned to obtain accurate estimates for income types and tax liability which are difficult to measure sub-annually. Experience in the development program has shown that respondents are willing to refer to tax returns to improve accuracy in classifying income sources and their levels.

1d. Noncash Income: Although many assistance programs provide noncash benefits and many jobs have significant fringe benefits, in-kind income has not been regularly or systematically measured in household surveys. The data which have been collected in earlier surveys mostly involve a few benefits received by the low income population [1]. A goal of this survey is to provide information on receipt (and changes in receipt) of selected in-kind income types. Since the estimation of the money equivalent of noncash benefits is a major issue, the survey will also collect information which can be used in valuation research.

1e. Tax Information. Household surveys that are currently available on a routine basis do not provide information on taxes. This makes estimation of tax incidence and simulation of policy alternatives in the tax area excessively time-consuming and imprecise. It also severely limits estimation of disposable income. An important objective of the new survey is to provide microdata on the amount and type of tax liability. This information will be collected in an "annual round-up" interview, and, at a minimum, will include income taxes, property taxes, and Social Security (FICA) taxes.

1f. Attitudes and Subjective (Self-Perceived) Well-Being. Data on the attitudes and perceptions of individuals are needed to understand participation in government programs, especially when lack of knowledge or fear of social stigma may act to deter program eligibles. These data also are essential to an understanding of the effect of changes in administrative practices and other aspects of program implementation on eligible recipients.

An important goal of the SIPP is the development, and subsequent routine collection, of improved measures of economic well-being. Researchers are increasingly aware that individuals' own perceptions about their economic situation offer a valuable, policy-relevant perspective from which to view economic well-being.

Since individuals' actions are based on their subjective assessment of economic realities as well as on the economic reality *per se*, subjective measures will be valuable input to microsimulation models. These measures can play a useful part in modeling a wide range of policy-relevant economic behavior, such as the timing of retirement, asset accumulation, and the intensity of work effort.

1g. Current Labor Force Participation. To meet its income data goals, the survey must obtain extensive labor force information. For example, many individuals receive income from more than one job, business, farm, or self-employment, and a large portion of program participants are receiving benefits because they are constrained to low wage work, are unable to find work, or are incapable of working. For those currently working in the market sector, information is needed about numbers and types of employment, hours and weeks worked, occupation, industry, and reasons for part-time employment. For those not working in the market, the areas of interest are the duration and reasons for unemployment or nonparticipation in the labor force, including the identification of discouraged workers and description of major activity or status (e.g., housekeeping, student, retired). These data are necessary for policy simulation and assessing program impact. Much of this information will be collected on a sub-annual basis, but the ability to calculate annual estimates will be maintained.

1h. Household Composition. The SIPP will collect extensive information on relationships within the household to approximate the key units of analysis. Analysis units will differ from program to program (such as OASDI, AFDC, SSI, Food Stamps) and for tax-filing units and other family groupings according to the policy purposes to be served. No single set of categories will be sufficient for all the analyses the SIPP data are intended to support. Moreover, since some members of the sample will change residence, marry, divorce, come of age, have children born to them, or die, the survey must be able to track them through their various changes and permit researchers to construct analysis units as needed. Consequently, the survey will collect and retain adequate data to permit construction of a variety of subgroupings within sample households.

1i. Personal Background. Data on experiences in each adult respondent's lifetime are needed to understand the personal circumstances that affect both program participation and earnings levels. A permanent feature of each SIPP panel will be questions on personal history. Among the most important areas are the following:

- Education and Training--Education is strongly related to earnings, and it is generally assumed that there is also a significant relationship between vocational training and earnings. The new survey will provide more detailed information on these relationships than has generally been available. In addition, there is interest in identifying those eligible for or participating in various federally sponsored educational programs. The developmental phase of the SIPP field-tested questions on courses taken in high school, highest college degree, major field of college study, vocational training programs, special schooling, and rehabilitation services.

- o Work Experience--Many studies of earnings differentials have concluded that prior work experience and on-the-job training account for a significant portion of differences in earnings levels and need for program assistance. This is potentially significant for evaluating such job-related programs as the Comprehensive Employment Training Act (CETA), Work Incentive Program (WIN), and "workfare" welfare reform proposals. The survey development program has tested questions on the number and length of interruptions in employment and the reasons for these interruptions (e.g., unemployment, health problems, and family responsibilities).
- o Migration--Information on place of birth and recent changes of residence will lead to a better understanding of past and current eligibility for public assistance programs as well as of current earnings. This information will permit improved analyses of such controversial issues as whether the "welfare" population moves to areas with higher benefits, or whether its members need assistance because they have not moved to areas where there are higher wages or lower levels of unemployment.
- o Marital History--Marriage, divorce, separation, and widowhood are frequently associated with major changes in income, family responsibilities, welfare eligibility, labor force participation, and program participation. Marital history data are therefore important for an understanding of current circumstances, and for learning more about the relationship between marital stability and program participation (especially in AFDC). Such data combined with work history information will allow the identification and study of displaced homemakers and other special groups.

1j. Variable Content Sections. An important feature of this survey is sufficient flexibility to include, on a regular basis, special questions or questionnaire modules on new subjects of particular interest. Topics of two general types will be included in these sections:

- o relatively stable topics that have considerable program and policy interest but that do not need to be collected every year or updated on each visit (e.g., wealth and pension coverage); and
- o emerging issues for which data are needed with short lead time (e.g., household responses to the 1979 energy crisis).

Among the topics examined during the developmental stage that are candidates for inclusion in operational panels are:

- o net worth
- o disability
- o health care and health insurance
- o inter- and intra-household income transfers
- o child care arrangements
- o perceptions of programs
- o educational enrollment and expenditures
- o life cycle financial planning and pension coverage
- o household energy consumption

2. Matched Data from Administrative Records. [2] The complete SIPP data base will consist of data gathered in household surveys, supplemented (as appropriate) by matched administrative data from program files such as the earnings and benefit records of the Social Security Administration or benefit records of the Supplemental Security Income program. Areas of current and growing concern are:

- o the potential impact of alternative Social Security benefit computation formulas on program expenditures and on the types and characteristics of families and individuals affected; and
- o the incidence of alternative FICA and general income tax formulas on the types and characteristics of families and individuals affected.

The long-term impact of differing benefit formulas is more a function of the earnings histories and retirement decisions of currently ineligible workers than of the characteristics of current recipients. To address this issue SIPP household and individual records will need to have Social Security covered earnings histories matched with the SIPP area sample cases. The developmental program experience indicates that Social Security numbers can be obtained for more than 90 percent of the covered sample. This match procedure is much more difficult than drawing list frame samples from, for example, SSI administrative records and matching SSI data, but is imperative for the development of an improved data base which will support analyses of the impacts of Social Security financing alternatives.

3. Design Features. The desired scope, timing, accuracy, and statistical reliability determine six major design features.

3a. Longitudinal Panel. The goals of measuring changes in income and economic well-being, identifying the causes or antecedents of change, and collecting extensive and detailed data on difficult-to-measure subjects require multiple interviews with the same individuals. These provide longitudinal depth and allow the recording of life events close to the point of occurrence. Measurement of change from period to period is more precise if the sample individuals are not changed. The longitudinal design also spreads the burden of response over several waves and reduces the length of each interview.

The longitudinal panel sample consists of the resident household members at the start of the panel. Each adult either will be interviewed or information about him will be obtained from a proxy respondent. A subset of the full interview will be obtained by proxy for children under 16 years of age. Sample persons who move singly or in groups will be followed in order to obtain subsequent interviews.

Interviews in the first panel will take place at staggered 4-month intervals over 28 months in order to collect monthly income and related data for two calendar years and tax information once each year. One year after the 28-month period, there will be a follow-up interview, thus expanding the longitudinal data files for special policy research purposes. A small new sample will be started in the second income-year to measure biases associated with conditioning and attrition in the main panel, and increase the sample size following sample losses during the initial year.

Interviewing for the first panel is to start in 1982 with 20,000-24,000 households; a mini-panel of 4000-8000 households will be added in 1983. Thereafter new panels will be proposed to provide annual measures of change over time in the major variables, especially income. As new panels are introduced, the income content and collection procedures will be kept sufficiently stable to achieve the income data and timely production goals of SIPP.

3b. Questionnaires. Questionnaires will vary from wave to wave in the SIPP panel. "Core" portions will change little; household composition and personal and family characteristics, labor force status and income profiles will be obtained in the first interview and updated in subsequent ones. "Topical" portions of the questionnaire will change from wave to wave, introducing new content areas which respond to emerging policy research needs. These may consist of a few questions or be comprehensive modules, according to the purpose being served. An annual "round-up" interview will collect tax liability and annual income estimates, based on W-2 forms and other records to the extent possible. Finally, the "year-later" questionnaire will be designed around special topics.

3c. National and Supplementary Samples. Most major surveys exclude the institutional population and part of the military from their universe and cover only the 50 States and the District of Columbia. The first SIPP panel will be similar. The program and policy usefulness of the SIPP's nationally representative sample can be increased by collecting policy relevant data from:

- o residents of the Commonwealth of Puerto Rico, who show very high rates of program participation,
- o persons in institutions (e.g., nursing homes, prisons, group residences), and
- o all members of the resident Armed Forces population.

The developmental program did not test the feasibility and cost effectiveness of alternative approaches to collecting these data, and such work must be carried out as developmental activities are continued in the 1980's.

As part of Census' sample redesign for the decade based on the 1980 Census, the sampling frame will be constructed to maximize the number of states for which state-reliable income and program data can be estimated from SIPP without seriously decreasing the reliability of national estimates.

The SIPP program also seeks to develop improved estimates for population subgroups such as high and low income groups and for larger minority groups such as Hispanics and Blacks. Census is investigating means to achieve better estimates as part of the sampling redesign for the 1980's, such as developing samples of minority addresses from the 1980 Census or other sources to supplement the national sample. Also, multiple frame sampling and estimation techniques are expected to increase the reliability of SIPP data for population subgroups such as program participants; for example, samples can be selected from administrative frames, especially SSI and OASDI. The possibility of using an IRS frame and several other frames (Aid to Families with Dependent Children, Unemployment Compensation, Veterans Administration, Workers Compensation) were tested during 1980 as part of the developmental program.

3d. The Use of Administrative Records. The use of administrative records will serve several purposes, namely to:

- o increase sampling efficiency for certain subpopulations (e.g., OASDI or SSI recipients),
- o supplement survey-reported data with administrative record data for items which are difficult to obtain in surveys (e.g., earnings and benefit histories),
- o provide more complete descriptions of the sample persons, and
- o allow comparisons between survey and administrative data for validation studies of specific items common to both data sources.

Records from HHS/SSA programs, and tax records from the Internal Revenue Service are primary candidates for consideration as sources of list frame samples. Supplemental frames initially will be treated as separate entities but some will be integrated into the national sample using multiple frame estimation techniques. Other studies involving evaluation of asset and pension data may, where possible, use respondent releases to collect and match data from other sources (e.g., banks and employers).

3e. Quality Control Program. The quality control program will consist of two major elements:

- o a current (or continuous) quality control system, and
- o periodic quality control projects designed to investigate special problem areas.

The continuous quality control system--where major features are outlined in a detailed management plan for data production--will monitor each stage of the survey process. Routine reports of data collection will allow close review and evaluation of noninterview rates, field office edit reports, observation reports, reinterview results, interviewer turnover rates, etc. Routine feedback from the edits will include counts of changes and written notification of other problem areas encountered by the clerical staff. In the computer processing area, tabulations by programers and analysts will show information such as edit failure rates by item, item nonresponse rates, coverage rates, results of weighting procedures, etc.

It is impossible to list target CV's for all subpopulations (e.g., for selected program participants and characteristics of program participants). However, based on some preliminary calculations, it appears that CV's of national estimates of 5 percent on items such as the 20th, 50th, and 95th percentiles of the family income distribution and the proportion of families receiving AFDC are both desirable and feasible. For these same variables, CV's of roughly 10 to 20 per cent for a number of states are desirable, given an emphasis on gaining some reliable state-specific estimates (even if it involves small sacrifices in national reliability estimates).

Oversampling and other techniques for reducing the variance of estimates for minorities and for the low income population need to be examined. Further study of additional issues related to the reliability of sample estimates under alternative sample designs is required before the design of the 1984 SIPP sample can be completed.

In addition, the SIPP has a commitment to investigate specific sources and magnitudes of the component of the variance and bias due to nonsampling error (e.g., between interviewer variability, response error, non-response bias, and bias associated with sample attrition). Implementation of a research plan for investigating these issues is planned.

Management And Access To Data

The Departments of Health and Human Services (HHS) and Commerce jointly sponsor SIPP. For HHS, the developmental program has been under the leadership of the Assistant Secretary for Planning and Evaluation (ASPE). The operational program will be under the leadership of the Office of Policy of the Social Security Administration through its Office of Research and Statistics (SSA/OP/ORS). For Commerce, leadership for the developmental and operational stages lies with the Bureau of the Census (BOC).

Within HHS, policy guidance is provided by the intradepartmental SIPP Steering Committee chaired by ASPE with membership from the principal operating agencies of the Department.

It is anticipated that SIPP data will be widely used by government agencies. Coordination and policy guidance at the interdepartmental level is provided by the SIPP Policy Advisory Committee chaired by the Office of Federal Statistical Policy and Standards (OFSPS) with membership from the Departments of Agriculture, Commerce, Education, Health and Human Services, Housing and Urban Development, Labor and Treasury, and the Council of Economic Advisors. (Participation by other agencies should be explored with the Policy Advisory Committee or directly with HHS/SIPP staff.) The Census Bureau and HHS will jointly develop and concur in operational commitments for the SIPP.

Also, SIPP data is expected to be widely used by private firms and the academic community. HHS and Census will issue periodic reports summarizing key results of the survey. In addition, both agencies agree to the desirability of timely release of fully documented microdata files. However, these files pose a special problem since an integral part of the survey plan is the inclusion of information from administrative data sources. Safeguards will be established which guarantee confidentiality (as defined in Title 13 or in new privacy legislation) while providing researchers with as much information as possible. Both standard public-use files containing no administrative data and special-use files containing a limited amount of administrative data may be released. The aim is to maximize the amount of data without compromising respondents' privacy.

Within confidentiality restrictions, data will be released to other governmental agencies. Currently data files from the research panels of the development program are being prepared without geographic identifiers or administrative data for public distribution as soon as adequate documentation can be made available to answer the basic information needs about the survey operations and microdata tapes. Census and HHS will continue to seek to accelerate preparation for timely public release of the data, and will issue public announcements at appropriate times.

Footnotes

[1] A frequent criticism of the March CPS income statistics (the official source of estimates of the poverty population) is that they exclude the value of noncash income items such as food stamps, free school lunches, publicly funded medical care, housing subsidies, employer provided health and pension benefits, etc.

[2] See also "3d. The Use of Administrative Records"