

**THE SURVEY OF INCOME AND  
PROGRAM PARTICIPATION**

**SIPP: FILLING DATA GAPS ON THE  
POVERTY AND SOCIAL WELFARE FRONTS**

No. 8705 *27*

**Paul Ryscavage**

**May 1987**

**U.S. Department of Commerce BUREAU OF THE CENSUS**

A version of this working paper will be published in a forthcoming book  
entitled The New Meaning of Poverty edited by Douglas Besharov and  
John C. Weicher.

## Table of Contents

Introduction.....	1
Existing Data Sources and Data Gaps.....	2
Data sources.....	2
Data gaps.....	3
An Integrated Approach to Data Collection.....	5
Policy Applications of SIPP.....	7
Poverty and income distribution.....	8
Program participation.....	9
Labor markets.....	9
Health.....	10
Education.....	10
Children.....	11
The family.....	11
Assets and liabilities.....	11
Fertility.....	12
Migration.....	12
The elderly.....	12
The SIPP Methodology.....	13
Sample design.....	13
Questionnaire design.....	14
SIPP data products.....	15
Summary and Conclusions.....	16
Notes.....	17

## INTRODUCTION

Every year, towards the end of summer, the Census Bureau releases its poverty statistics for the preceding calendar year. These data are derived from the Current Population Survey (CPS), a survey of approximately 60,000 households. Frequently, the Bureau's figures are disputed because of how poverty is defined and how it is measured in the CPS. As most observers of the poverty and social welfare fronts know, the Federal government "statistically" defined poverty and began measuring it back in the 1960's. Despite their controversial nature, however, the Census Bureau's poverty data have been used by policy makers for setting the government's poverty and social welfare policy for many years.

We have also learned much about the poor and the nature of poverty from the Census Bureau's estimates. We know who the poor are, where they live, what income sources they have, the extent to which they participate in the labor market, and so on. We also know how well we have done in combatting poverty (according to the official definition) over the last 25 years or so. But, as many observers have pointed out, the existing statistics from the CPS don't provide all the answers.

The problem of poverty that confronts our Nation today is different than it was in 1960. In many respects, the problem has become more complex. One unanswered question is whether or not in our efforts to alleviate poverty, we have only made it worse. Another concerns the importance of the social and cultural dimensions of poverty vis-a-vis the economic dimension. Like many diseases, poverty may develop an immune system which seals itself off from traditional cures. And specific gaps in our poverty statistics have also been identified. For example, we need to know more about the impact of noncash benefits (i.e., food stamps, subsidized housing, Medicaid) on the poor; we need to know more about the changing household composition of the poor over time and the changes taking place in their income flows; we need to know more about how the labor market and economic hardship are interrelated and whether or not the labor market is failing the poor or vice versa.

To provide us with the answers to many of these questions, we need a data base that captures the "dynamic" nature of poverty, that is, a data base that is more reflective of the poverty process itself. Indeed, we need a new microscope, a more powerful one, by which to examine this Nation's poverty problem.

In October of 1983, the Census Bureau began a new household survey called the Survey of Income and Program Participation (SIPP). The survey was designed to provide comprehensive information on the economic situation of households and persons in the United States. The survey has two distinguishing features: it is an integrated survey designed to collect data regularly on a wide variety of topics relating to this country's well-being; and it is a longitudinal survey in which the same persons are interviewed every four months over a two and one-half year period making it possible to observe changes in well-being. Because of its design and scope, SIPP stands at the forefront of survey technology.

In the next several pages we will discuss the CPS and other poverty data sources that have been used over the years and upon which has been built much of our poverty and social welfare policy. Data gaps are then discussed. The SIPP is then presented as a integrated data source which not only addresses the

data deficiencies on the poverty front, but which may also provide important insights into other social issues such as youth unemployment, health insurance, and the family. Potential policy applications of SIPP data are then discussed in detail. The chapter closes with a description of the SIPP methodology and the content of the survey and its potential products.

The SIPP is in its infancy relative to the other surveys that have provided poverty data and social welfare statistics. Its sample size is still relatively small, the survey design and methodology are new, the processing and estimation procedures have just recently been set in place, and interviewers and field personnel have had to become familiar with a very complex questionnaire--in short, SIPP is in a "settling-in" process. During this process, changes in all aspects of the survey are possible with a resulting impact on the data. Nevertheless, and as with all new technologies, these refinements will only enhance the instrument.

#### Existing Data Sources and Data Gaps

Since the compelling expressions of concern about the Nation's poverty problem over 25 years ago, policy makers and researchers have wanted to know more about the poor. To this end, they turned to the Nation's survey takers and information collectors. Together they modified existing surveys, began new ones, and, in general, searched intensively for data about the poor. What resulted was a cornucopia of statistics of different kinds and qualities. The following discusses some of the more well-known sources relied on by policy makers and researchers in the past.

Data Sources. The CPS, as mentioned at the outset, has been and continues to be the principal source of information on the poor. This survey's primary purpose, however, is to collect monthly data on the Nation's employment and unemployment situation. 1/ The poverty data are obtained each March through a supplement to the monthly labor force questions. This supplement consists of a battery of questions relating to the previous year's money income and sources. A few questions are asked about nonmoney sources of income, such as food stamps, Medicare, and employer-provided health benefits, but no value is placed upon these items and they are not included in the Census Bureau's income estimates.

From these estimates of cash income received by the Nation's households, the official poverty definition is applied and the poverty population determined. 2/

Many of the existing data deficiencies in the poverty statistics from the CPS have emerged because of the way in which the income data are collected. That is, the income questions were "tacked on" to a labor force survey. The original design and methodology of the CPS was aimed at collecting monthly data on the number of employed and unemployed persons. And this was because the CPS was developed at the end of the Great Depression when national concern was focused on mass unemployment. The idea to collect income data through the CPS was an "after thought" so to speak--and regular collection did not begin until the end of the 1940's. As a matter of fact, other supplements dealing with a variety of topics, such as fertility and education, have been added to the employment and unemployment questions down through the years.

## Table of Contents

Introduction.....	1
Existing Data Sources and Data Gaps.....	2
Data sources.....	2
Data gaps.....	3
An Integrated Approach to Data Collection.....	5
Policy Applications of SIPP.....	7
Poverty and income distribution.....	8
Program participation.....	9
Labor markets.....	9
Health.....	10
Education.....	10
Children.....	11
The family.....	11
Assets and liabilities.....	11
Fertility.....	12
Migration.....	12
The elderly.....	12
The SIPP Methodology.....	13
Sample design.....	13
Questionnaire design.....	14
SIPP data products.....	15
Summary and Conclusions.....	16
Notes.....	17

Another important source of poverty data which has yielded many insights into the problem has been the Panel Survey of Income Dynamics (PSID). 3/ Begun in the late 1960's, the PSID is a longitudinal survey which started operation with approximately 5,000 households in its sample. These households were interviewed every year concerning their incomes and other social and economic developments in the preceding year.

One of the major contributions of the PSID was the discovery that longitudinal data--data based on the same individuals or families over time--yield a much different picture of the poor than the cross-sectional data from the CPS--data based on different individuals and families at different points in time. From the PSID it was learned that the individuals living in poverty one year are very likely not to be in poverty the next--in other words, a lot of turnover exists in the poverty population. Moreover, the PSID data also revealed that changes in households' economic situations are closely related to changes in household composition, whether because of separation, divorce, marriage, death, etc. 4/

Although this data source has been very important to our understanding of poverty, it too has deficiencies which limits its usefulness. One problem is that its sample is very small and it becomes difficult to make inferences about small segments of the population, for example, black youths from inner cities. Another problem is that the data are collected on an annual basis only, making it impossible to observe income changes within the year.

Two other sources of data that have been used extensively are the Survey of Economic Opportunity (SEO) conducted in the mid-1960's and the Survey of Income and Education (SIE) conducted in the mid-1970's. Both of these surveys were "one-time" surveys of the population, but yielded important insights into the nature of poverty at those times. The SEO was a large household survey in which black households were oversampled in both 1967 and 1968; the SIE was an even larger survey designed to make poverty estimates by state and collect information about education levels of children in 1975.

Data gaps. Researchers for many years have pointed out the data gaps in our present day poverty statistics. Some of these relate to the way income statistics are collected presently in the CPS, while others relate more generally to deficiencies in survey methodology.

First, there is a need for better data on "noncash" income received not only by the poor, but the nonpoor as well. As most income analysts know, our incomes today are composed not only of money incomes but nonmoney incomes as well.

The classic examples of nonmoney income are Food Stamps, Medicare, and employer-provided health insurance, but of course there are many more. Beginning in March 1980 information about the reciency of certain noncash income types was collected in the CPS and the Bureau of the Census has regularly reported on them. 5/ Of equal importance to their collection is the valuation of these various noncash income items; once they have been valued, they can be included with the cash or money income estimates. But valuation is problematic. Survey respondents, in most instances, have little idea as to the value of specific noncash benefits. As the Bureau of the Census has shown, however, inclusion of the value of noncash income has a significant effect on the measured size of the poverty population and the shape of the income distribution. 6/

A second deficiency in the income statistics from both the CPS and PSID is that they are collected annually. Collecting income data on a subannual basis would be preferable for two reasons: First, respondents, and, in particular, respondents who are poor, have income flows which vary during the year and second, it is easier to remember incomes received over short periods versus 12 or more months. Income flows could be monitored better and income reporting would be more complete.

Third, many researchers have suggested that only with a large longitudinal survey conducted during the year over a number of years can we observe the dynamic nature of poverty. That is, they need the capability to follow the same persons through time to observe changes in their economic situation. Only then can correlates of economic change be identified, and the appropriate policies formulated.

With the exception of the PSID, data sources in the past only provided us with cross-sectional views, or snap shots, of the poverty population. We have missed the gross flows, or the movements of households into and out of poverty.

As the PSID has suggested, the poverty population is constantly changing. What's needed, however, is a closer look, a more focused look, at the turnover in this population.

Longitudinal data also provide other insights that cross-sectional data cannot. For example, how long, on average, is a completed spell of unemployment? With longitudinal data this can be calculated but with cross-sectional data we must be satisfied with a synthetic estimate or an estimate based on a long recall period. Similarly, it becomes possible with longitudinal data to calculate spells of food stamp reciprocity or some other spell of Federal government program participation. In addition, longitudinal data enable one to see the changes in particular statuses, for example, the movement of individuals or households into or out of a specific transfer income program and the associated changes.

A fourth gap in our poverty statistics is that it is difficult to determine who is and who is not eligible for Federal assistance. This is because in the CPS, income is measured for a year but for many programs eligibility is based on monthly income. This deficiency is particularly serious for those policy analysts building micro-simulation models from which policy decisions concerning eligible universes are to be estimated. Also, it is difficult to determine how well particular programs are targeted (i.e., what proportion of the eligible population is participating in the program?)

Related to the previous data gap is a fifth deficiency. Past data sources have had little information on assets and wealth. Certain Federal income transfer programs require information on a household's asset position before eligibility is decided. More importantly, it is essential to have asset information for assessing the economic well-being of the poor and to what extent assets (e.g., owner occupied homes) are relied upon as sources of income.

A sixth gap involves the relationship between the labor market and economic hardship. While information is available from the CPS and PSID on this relationship, in both surveys it is based on a single interview reviewing experiences which took place in the past year or so. It becomes difficult to "link" problems



in the labor market (e.g., unemployment, low wages, economic part-time) to economic problems in the household. What is needed is an accounting of labor market experiences and income changes over a short period of time.

A seventh and last data deficiency is specific to the CPS. Households are interviewed every March about their money incomes received in the previous calendar year. The incomes of persons who have moved out of households at the end of the year would not be reported at the time of the interview, even though these incomes were used to sustain the economic health of the households. Conversely persons who join a new household at the end of the year and are asked about their incomes would have them reported with the members of the households they just joined. Obviously, this inconsistency in income reporting causes problems for poverty measurement.

In short, the deficiencies of our present day poverty statistics can be boiled down to three.

1. They don't reflect income from all sources nor the full economic situation of households.
2. They don't reflect the dynamic nature of poverty, or the movement of households into and out of poverty.
3. They don't reflect the interrelationships existing between the labor market, the Federal government, and poor households.

Closing these data gaps is essential if we expect to develop effective poverty policies in the years ahead.

#### An Integrated Approach to Data Collection

Federal government statistical surveys have typically been subject--specific, that is, if it was decided information was needed on crime a crime survey was taken, or on health, a health survey was begun. As mentioned earlier, the CPS was established in the early 1940's for the purpose of measuring unemployment and the SEO was conducted in the mid-1960's to learn more about the poor. While this approach is natural and logical, it may not be the ideal approach when the subject of concern is more general in nature, such as economic well-being. This is because a topic like well-being is affected by so many things: jobs, housing, income, health, crime, and so on.

Policy makers who attempt to assess the economic well-being of Americans have had to assemble data from a variety of sources before making judgements and reaching decisions. And when new information was needed on a particular subject, they frequently added supplemental questions to ongoing surveys. Consequently, some surveys which began with limited data goals evolved slowly into multi-propose surveys. The classic example, of course, is the CPS -- the source both the "official" poverty statistics and "official" unemployment statistics.

All of the CPS supplements yielded--and continue to yield--important information about aspects of well-being of the population. Unfortunately, the supplemental information collected in one month cannot be easily related to data collected in another month's supplement. For example, it is not possible to examine the

fertility patterns of poor households because fertility questions are asked in the June survey while income questions are asked in March.

Because of problems like this, researchers and policy makers cannot "integrate" the different types of data collected in the CPS that are policy relevant. The sample design prevents this. In other words, the evolution of the CPS was constrained by the initial purpose of the survey which was to collect mostly unemployment information.

The Survey of Income and Program Participation (SIPP) began in response to many of the data deficiencies listed earlier and, in general, to a dissatisfaction with the precision and scope of the income and program data derived from the March CPS. Consequently, the SIPP was designed differently than the CPS and with a much broader purpose in mind. While its title suggests that it primarily collects data on income and persons participating in transfer income programs, it is safe to say that SIPP has many data purposes, all of which ultimately relate to the well-being of persons and households. In recent years our society has become considerably more complex both in terms of where our incomes come from and in our relationship to the government. Our well-being today does not simply depend on whether or not we have a job, for example, but whether or not it is a good job or whether or not we qualify for food stamps or whether or not our employer pays for part of our health insurance. Consequently, the SIPP is aimed at collecting many data elements which go into determining our society's well-being.

But not only does the complexity of our society require an integrated approach to data collection, other factors demand it as well. The SIPP reflects perhaps the two most important influences on our household's well-being, the labor market and the government. Our social welfare system has so evolved that if the labor market does not provide an acceptable means of support, the government will step in (i.e., "the safety net"). The SIPP provides policy analysts with the opportunity to see how this relationship is faring.

The policy issues of today will continue to change and new ones will emerge. This is the nature of any modern and dynamic society. Unlike the CPS in which new data requests were added to a labor force survey, the SIPP permits a more natural evolution. This is because the basic components of well-being--labor force activity, income, program participation--are already in place at the heart of the survey. Data demands arising from policy makers can be accommodated in the "topical modules" (similar in concept to the CPS supplements) of SIPP and can be related to other topical modules as well as the core information.

Social scientists today realize that to answer questions concerning behavior in our Nation we need longitudinal information. Processes must be better understood: What are the events that push a family into poverty? Do our social welfare programs help people climb out of poverty or do they perpetuate poverty? Is the labor market failing certain groups of people or are they failing the labor market? The SIPP is collecting the information necessary to provide some of the answers to questions like these.

In summary, SIPP was designed to longitudinally collect subannual data on

- -- money income
- -- noncash income
- -- labor force activity
- -- household composition
- -- program participation
- -- eligibility for program participation
- -- characteristics of program participants

and collect personal characteristics relating to

- -- taxes
- -- assets and liabilities
- -- work and education histories
- -- health, disability, and life insurance coverage
- -- migration
- -- past fertility

in order to produce policy relevant statistics on individuals' and households'

- -- spells, durations, and frequencies of economic hardship;
- -- transitions in labor force, program participation, and economic status;
- -- economic resources.

In meeting the objective of an integrated data source, SIPP will provide new ways of observing and understanding the Nation's poverty problem and economic and social well being. Policy makers will have new ways of posing policy questions, and analysts will have better information to answer those questions. The policy community will be able to understand not just events and characteristics, but also the sequencing of events and patterns of activity that will allow more innovative solutions to social problems.

#### Policy Applications of SIPP

Policy analysts and policy makers have traditionally used data in their decision making process in three ways: basic estimation, modelling, and micro-simulation.

By basic estimation we mean simply, using the survey as the source of a data estimate. For example, how many black families slipped into poverty during 1987? Or how long was the average spell of AFDC reciprocity? Or how many households were participating in more than one means-tested cash or noncash transfer income program? Here the policy analyst is in need of reliable, up to date information, as soon as possible.

SIPP can also serve the needs of "model builders," whether they be in the policy arena or in the research community. The point of modelling is to learn about or understand a particular behavior. In the context of household formation, for example, policy makers would be interested in knowing more about the reasons for its fluctuations. The data are used for more than just descriptive purposes: they are used to help explain behavior.

The last way in which SIPP data can be used by policy makers is in micro-simulation. Behavioral processes are simulated using micro-level data (e.g., persons, households) and, on occasion, the behavioral parameters estimated in models. These techniques have been relied on extensively for examining the ramifications of changes in such things as tax laws and program participation rules. Microsimulation is an important policy tool and one of its requirements is a data base that is sufficiently large with many subject areas.

Because the SIPP data base is derived from a representative sample of households around the country, its policy relevance extends well beyond the problem of poverty. Indeed, SIPP data have relevance for a variety of social welfare issues. In the following discussion we focus on how the SIPP data might be used in addressing poverty and social welfare topics.

Poverty and income distribution. Perhaps the single most debated issue involving the income distribution concerns poverty and income inequality. Our existing statistics indicate that while some Americans live very affluent lives, others are very poor.

The primary source of income inequality in our society is the labor market and the distribution of earnings it generates. Some of the observed inequality is due simply to the rise and fall of earnings over the life cycle: younger and older persons earn less than those individuals in the middle years. Consequently, if income inequality is to be dealt with through public policy, the policy decisions must focus on the "other" causes of earnings inequality. The other causes of inequality reflect differences in individuals' abilities to contribute to national output whether because of inherited aptitudes or environmental influences. Here, specific policy choices for redistributing income emerge.

While policy makers know how many poor persons there are and their characteristics, they need to know more about the dynamics of the poverty population. For example, high on their list is information regarding the rates of escape from and entry into poverty.<sup>7/</sup> Entry and exit rates determine the poverty population at a point in time. Depending on what these rates are--and their correlates--will help determine the kinds of policies that are adopted. This kind of information can be obtained through SIPP because of its longitudinal design.

Another application of SIPP data involving poverty policy is in the area of microsimulations. As mentioned earlier, microsimulation techniques require large data bases reflecting behavioral responses of individuals. For example, what would the labor supply response be among the poor if day care centers were available? Or what would be the effect on the size of the poverty population of a massive welfare reform initiative? Answers to hypothetical questions like these could be derived using the SIPP data in a microsimulation model.

The Federal tax system has been long considered as one of the means for redistributing income in this country. Some have claimed, however, that the overall national tax structure (Federal, State, and local taxes) has very little effect on the distribution of income.<sup>8/</sup> Whether it does or doesn't, the SIPP collects data on Federal and State income taxes, as well as local taxes. Other tax information is collected such as the type of return, use of selected

schedules, and number of exemptions. Information is also obtained on property taxes. Not only will this tax data help in better estimating the after-tax income distribution, but it will also be useful, along with information from the income transfer system, in refining microsimulation models of the tax and income transfer system.

One other use (but certainly not the last) of the SIPP for income distribution analysis is in providing more information on "noncash" income. As is well known, the income concept currently used by the Census Bureau is a cash income concept. In recent years, efforts have been made to value many of the more common noncash payments (e.g., food stamps, Medicare, Medicaid, certain employer-provided benefits) because of their growth in importance to households. Indeed, one economist has said: "The issue of measuring and valuing nonmoney income has finally come into the mainstream of economic policy analysis."<sup>9/</sup> SIPP provides an abundance of data on noncash income: food stamps, Medicare, Medicaid, subsidized housing, free or reduced-price school lunches, energy assistance, employer-provided pensions, health insurance, private disability insurance, maternity leave, employer education and training benefits, sick pay, and so on. While the valuation of many of these items is no easy task, the basic information will have been collected.

Program participation. The number of transfer income programs in existence is not only large but also exceedingly varied. Many of them are aimed at the poor and low income groups, while others are entitlement programs available to persons across the income distribution. SIPP can monitor them all. It should be able to tell policy makers who should and should not be in specific programs as well as how effective the programs are.

One very topical question raised by Charles Murray in his book Losing Ground is whether or not the means-tested programs foster a sense of dependency which ultimately affects decisions about work and the family. The SIPP data base should be able to speak to this issue. A second question involves the long term effects of program participation. That is, does participation in a program improve the chances of success in the job market or in school or in the family? Are the households that move out of poverty staying out of poverty? A third question involves whether or not programs overlap one another. Some answers have appeared in the literature <sup>10/</sup> and the SIPP should be the chief data source used to address this matter in the future.

Labor markets. Exemplifying the integrated nature of the SIPP are the battery of questions related to labor force participation. Work continues to be the chief source of income for Americans and is an obvious dimension in the country's well-being picture. People who have difficulties in the labor market are of concern because the difficulties often produce economic hardships. For many years economists have tried to measure the hardships caused by unemployment, low wages, involuntary part-time employment, and so on, but with only limited success. <sup>11/</sup> To fully understand the relationship between a labor market problem and its impact on the household requires longitudinal data. SIPP should be able to tell us about the impact of labor market problems.

Related to labor market-related economic hardship is the information SIPP provides about changes in labor force status from month-to-month, or transi-

tions, and how they are related to a household's economic situation. The labor market is a very dynamic place with millions of people flowing in and out, as well as between employment and unemployment. One transition that labor economists have been particularly interested in for many years is the one occurring after a spell of unemployment has been completed, that is, did the person find a job or drop out of the labor force. Using CPS gross flow data, Clark and Summers estimated that 45 to 50 percent of all unemployment spells end by labor force withdrawal. <sup>12/</sup> With SIPP it will be possible to identify job terminations, observe the spells of unemployment, and determine not only its duration and outcome, but the impact of it on the household.

Other policy applications of the SIPP labor force data involve occupational and industrial mobility. With this data we should be able to assess the changing infrastructure of our economy. Data relating to the "reservation wage" (or lowest wage for which one will work) should help us assess the jobseeking intentions of the unemployed.

Health. SIPP will provide important data for policy analysts in the health area. Specifically, the data will address issues of health insurance coverage and health care utilization.

Each wave of interviews contains questions designed to find out whether an individual has Medicare, Medicaid, and/or private health insurance. It is also possible to determine if a person had a health insurance plan provided by an employer or union, and if so, whether they paid part or all of the costs, and whether it was an individual or family plan. Unlike some health surveys, SIPP distinguishes between specific children in the household and insurance coverage within subfamilies of the household, which is particularly important in poor households where multiple subfamilies may be present. <sup>13/</sup>

Because of the longitudinal nature of the SIPP, it will be possible to track the stability (instability) of Medicaid coverage as changes occur in employment and family composition. Characteristics of the uninsured population will be available as well as characteristics of the insured who have overlapping insurance coverage. And most importantly, the longitudinal feature of the SIPP allows analysts to observe the relationship between private health insurance coverage and spells of short-term and long-term unemployment.

Education. Education figures prominently in policies designed to reduce poverty, as it does in the Nation's attempt to remain economically competitive in world markets. Moreover, education fosters opportunity and social mobility.

Because of its importance and the substantial investment our society makes in it, it is essential that educational policies be based on a sound understanding of the "causes and consequences" of educational attainment. Perhaps the chief policy use of the SIPP data on education will be to assess the relationship between education and training on the one hand and labor market participation and income on the other. <sup>14/</sup> Structural unemployment, the shifting economic base of the economy, and the popular concern over the quality of education in America demand this assessment.

More must be learned about the school-to-work transition and this is particularly so for those who do not go on to college or even finish high school. SIPP's longitudinal dimension will monitor that process. A significant amount

of data is collected on educational assistance provided by Federally funded programs as well as educational financing. Moreover, educational histories are collected in the SIPP which can be easily related to one's current economic circumstances.

Children. In recent years the Nation has witnessed an increase in the number of children living in households that are poor. Related to this has been the decline in the proportion of children living in two parent families, especially among blacks. SIPP data should be useful in addressing this issue.

The longitudinal dimension in SIPP offers analysts a "close-up" view of the economic and social environments of children. <sup>15/</sup> Events taking place in the household (e.g., a separation, death, spell of unemployment) very often have direct consequences on the welfare of children. From SIPP it may be possible to learn more about the factors responsible for the increase in families with female householders. For example, how important is teenage pregnancy or joblessness among young men? Data will also be available to tell us how effective child support payments are in maintaining families in which the father is absent.

Another set of data related to children is concerned with child care. Questions about child care arrangements, such as who provides the care, the number of hours a week child care is available, where the care is provided and its cost, should elicit useful information for policy discussions. Furthermore, data from the welfare history questions will measure the extent to which children have been dependent upon government transfer programs in the past.

The family. Concern over the future of the family has also become a topic of much debate in recent years. Dramatic changes in the rates of marriage and divorce have prompted some observers to predict the end of the family. Does the increase in persons living alone or with other unmarried individuals mean that our society is slowly evolving new forms of "family" relationships?

The SIPP survey design enables researchers and policy analysts to explore the "joint process" of household change and economic change. <sup>16/</sup> It is possible in the SIPP to observe such changes on a monthly basis over a 2 1/2 year period. In the survey, original household members are followed to new households permitting the mapping of extended family arrangements.

Marital dynamics, of course, are related to other behavioral phenomena such as fertility, geographic mobility, and labor market activity. Consequently, it will be possible to observe the impact of a separation or divorce on other variables which are present in the SIPP. The marital history data will yield important dates by which to calculate duration measures that can then be related to other reported behavioral changes.

Assets and liabilities. One of the major features of the SIPP data base is the collection of data on assets and liabilities of individuals and their households. <sup>17/</sup> Previous to SIPP only limited data have been collected from relatively small samples.

The asset information consists of such items as the current value of one's house, the value of vehicles and life insurance policies, the amount of money

in checking accounts, the amounts of stocks and bonds and other financial assets. Information on liabilities comprise such items as amounts owed on mortgages, doctor bills, margin accounts, and so forth.

A number of policy areas can be examined with such data. One of these concerns programs eligibility. For example, how many persons would be affected by a change in asset requirements for a certain program? Another concerns the elderly and the relationship between health, wealth, and retirement. Do the elderly begin to use their savings and convert assets to live on when they retire or do they continue to save? How liquid are the assets of the elderly? Do those nearing retirement accelerate mortgage payments to own their own homes when they retire? A third area concerns the young. What is the saving rate of the baby boom generation and, in general, their asset accumulation pattern? Do the SIPP wealth data reflect the "hard times" that the media say young people are having?

Fertility. Early childbearing often results in dependency on public programs for income support because young mothers are not prepared to support themselves and the fathers may or may not be able to support them. Some of the SIPP questions on fertility concern adolescent childbearing.

SIPP measures the variables necessary for determining the economic well-being of teens and their children, that is, their sources of income, the resources of their parents (if living with the parents), and the programs the young mothers may be participating in. In addition, it becomes possible to track these mothers experiences with the labor market and school system. One would also be able to observe the circumstances surrounding teenage childbearing vis a vis childbearing among an older group of women (say 20 to 24 years old) and compare differences.

Migration. SIPP sample persons who leave a household are subsequently followed for additional interviews (according to certain following rules) making it possible to study migration and the migration process. A number of policy-related issues might be addressed. First, the SIPP permits one to search for the correlates of an individual's migration decision whether it be a lost job, a divorce, a death, or some other factor associated with the household. Second, the SIPP data might reveal whether migration decisions were based on opportunities for program participation. And third, the SIPP may enable one to determine whether migration was due to seasonal factors (migrant workers) or life style choices (retirees). Related to this, of course, is the subject of migration and economic conditions. Do more people migrate in good times or in times when the economy is slumping?

The elderly. The SIPP will be an important data source for policies involving the elderly. The survey may eventually serve as a basis for studies of differential mortality relating to socioeconomic status, marital status, and other economic categories. The effects of a death on the living arrangements of the elderly will be recorded. That is, it will be possible to observe the effects on household finances, household program participation, and so on both before and after a death has occurred. Also, it may tell us the extent to which the elderly supplement their Medicare and Medicaid coverage with private health insurance. For many elderly persons private health insurance has assumed great importance. And last, valuable data are collected on who cares for the



ill and disabled in a community. Questions on long term care are important for assessing as well as for planning health care needs.

### The SIPP Methodology

The genesis of the SIPP was the Income Survey Development Program (ISDP) initiated by the Department of Health, Education, and Welfare in 1975. <sup>18/</sup> The purpose of the ISDP was to research ways in which the Federal government could improve its collection of cash and noncash income data and to begin the collection of data related to the various transfer income programs operated by the government. Between 1975 and 1981, new procedures were developed and tested for collecting this type of information in a household survey on a subannual basis and longitudinally. Much of the work centered around four experimental field tests that were conducted in collaboration with the Census Bureau. Based on the work of the ISDP, planning began for implementing a new survey called the SIPP. In the fall of 1981, however, funding for the research and planning leading to SIPP was deleted from the budget of the Department of Health and Human Services (formerly Health, Education, and Welfare). Work on the SIPP came to a halt. By 1983 new funding was appropriated in the Census Bureau's budget for the SIPP.

Sample design. Since one of the original goals of the ISDP was to improve the reporting of income and program-related information so that one could analyze changes on a subannual basis, the sample design had to accommodate the collection of a large variety of information, some of which would be collected over and over again. These goals were met principally by using a design in which the same people would be interviewed more than once, in other words, a longitudinal survey.

Persons in the SIPP sample are interviewed every four months for approximately 2 1/2 years. Sample persons are interviewed at new addresses if they move, and any other persons that they move in with are also interviewed. In this way, a highly detailed record of information is built up over time for each person and household in the sample. The reference period is the four months preceding the interview. For example, in October 1983, the reference period was the June-to-September period, and in February 1984, the time of the second interview, the reference period was the October-to-January period.

The first SIPP panel, designated as the 1984 panel but begun in October 1983, comprised 20,000 interviewed households. To create manageable interviewing and processing work loads, sample households were divided into four subsamples of nearly equal size. These subsamples are called rotation groups, and one rotation group, or one-fourth of the sample, is interviewed each month. Thus, it takes four consecutive months to interview the whole sample, and this four month period of interviewing is called a wave.

As shown in Figure 1, in February 1985 a second panel--the 1985 Panel--was introduced and it comprised 14,000 households. A third panel was started in February 1986 consisting of 12,500 households and panels of about this size or larger are expected to be fielded every year. Obviously, this "panel" design allows the overall sample to grow (or contract) through time as well as to change in content as the data needs and requirements change. This design feature permits a more natural evolution in the survey content than has been possible in other household surveys.

Questionnaire design. The SIPP questionnaire is long and complicated, but trained interviewers can administer it with ease. The interview with each household respondent age 15 and over is planned to last about 30 minutes. The content of the questionnaire is divided into three parts. The substance of two of these is essentially the same for each wave of each panel, while the third changes in each wave of a panel.

The first part of the questionnaire is referred to as the "control card." This document is actually separate from the questionnaire. It lists every individual at an address and records their age, sex, race, educational attainment, marital status, and so forth. The card is updated at subsequent interviews to record any changes that have taken place in the household, for example, the birth or death of a household member. During each interview, each source of income is listed as well as the name of each job or business.

The second part of the questionnaire consists of the "core" questions and is divided into five sections. The first section consists of questions pertaining to labor force activity in the reference period and the receipt of income from various sources; the second section obtains information about the receipt of wage and salary earnings; the third section focuses on earnings from self-employment; the fourth section collects the amounts of money received from the income sources mentioned in section one; and the fifth section collects amounts of money earned on asset holdings, such as savings accounts, bonds, stocks, and rental property.

The third part of the questionnaire--the topical modules--consists of supplemental questions related to the information collected in the first two parts. These topical modules are intended to meet the survey's goal of flexibility. There are two types of modules, fixed and variable. Fixed modules are generally the same from panel to panel. The variable topical modules are designed to satisfy the special needs of Federal agencies. A wide variety of topics are addressed in the topical modules and they give SIPP its "holistic" approach to well-being. (A review of the subject content is contained in the next section.)

In summary, there are three major characteristics of SIPP's methodology: First, SIPP's longitudinal design permits the monitoring of changes in well-being for a 2 1/2 year period and also allows researchers to examine the correlates of observed changes in well-being; second, the core-module design of the questionnaire permits a wider view of the many factors affecting well-being than other household surveys, while at the same time maintaining a focus and net profit or loss of the firm, about the number of people they employed, and so on.

Questions about program participation are asked. In addition to finding out about participation in Social Security, Medicare, Medicaid, Supplemental Security Income program (SSI), the Veteran's Administration pension program, the Food Stamp program, and others, the questionnaire also inquires about subsidized housing, energy assistance, and free or reduced-price school lunches and breakfasts.

The content of the topical modules is varied. Figure 3 presents a summary of their content in the 1984, 1985, and 1986 panels. Many of the same subjects are asked about from panel to panel as well as within the same panel. This allows analysts to not only observe changes in certain variables over time, but also to observe how these items change in a particular family or for a group of individuals.

The topics are all concerned with various aspects of well-being. For example, in wave 3 of the 1984 panel detailed information is asked about health and disability, work history, and education history. In waves 4, 6, 7, and 9 the focus is on economic aspects of well-being, that is, assets and liabilities, annual income and taxes, employee benefits, and so on. In wave 8, the fixed topical module focuses on personal histories, that is, fertility, marital, and migration histories. The variable portions of the modules in the 1984 panel examine topics of immediate policy interest, such as child care arrangements, welfare histories, disability status of children, and home financing arrangements.

One significant difference between the contents of the three panels is that in the 1986 panel all of the history-type questions were asked in the second wave. This was done so that critical events in individuals' and households' lives occurring before the survey could be dated making events taking place while they are in the sample more meaningful.

SIPP data products. A variety of data products will emerge from the SIPP. Most data users are accustomed to dealing with data that relate to some type of unit at a point in time, for example, the number of persons unemployed in February 1987. SIPP data products will not only include this type, but others as well.

Figure 4 presents a typology of the potential data products that can be generated in SIPP. <sup>20/</sup> The two dimensions of the typology are TYPE OF ANALYSIS and UNIT OF ANALYSIS. Because of SIPP's survey design, two types of analysis are possible, cross-sectional and longitudinal. Units of analysis can be observed at a point in time (a snapshot) or through time (a panorama). Within the cross-sectional analysis we can examine a unit of analysis at a point in time (statics analysis) or at two points in time (comparative statics). Within the longitudinal analysis we can observe a unit of analysis at the beginning and end of a period (also comparative statics) or observe the whole period (dynamic).

The unit of analysis may be persons, families, households and even a period of time, or spell. A spell is time-dimensioned in that it is a period of time a person, family, or household might experience something. For example, in the course of a year, one person may have three spells of unemployment or a family one spell of being on food stamps.

Cells 1 and 3 of the figure are the familiar data products that are generated in the CPS. For example, cell 1 might be the number of persons receiving food stamps in February 1987 and cell 3 would contain a comparison of such persons in February 1987 and February 1986. Cell 5 contains a product in which the characteristics of the same households, families, or individuals are examined at the beginning and ending of a specific period of time. An illustration might be the following: Of the individuals who were married on January 1, how many were still married on December 31? This type of analysis is frequently referred to as "gross flow" analysis and can be carried out with the SIPP data.

Data products in cells 2, 4, and 6 parallel the products in cells 1, 3, and 5. The difference is, however, the unit of analysis is a "spell." An example of a product in cell 2 would be a three month period of unemployment or food stamp

reciprocity. Cell 4 would contain products in which the number of spells are compared for two periods of time (e.g., there were 15 million spells of unemployment in year X and 18 million in year Y). Cell 6 involves changes or "transitions" in spells. An example of a product here might be the characteristics "before and after" a transition (e.g., family income before a marital separation and after a separation).

The data products in cells 7 and 8 are involved with actual processes, or "dynamics." In cell 7, for example, the product would reflect the patterns experienced when moving from the beginning of the period to the end of it: a worker's labor force status pattern (e.g., employment and unemployment) in a year. The products in cell 8 would attempt to identify the significant factors associated with an occurrence of a transition or change in spell. In this category, data products become "causal" in nature because they are attempting to explain why behavior has changed.

### Summary and Conclusions

The SIPP will not only fill many of the data gaps on the poverty and social welfare fronts, but it will provide policy makers with new insights into the ways in which American households function and respond to change. To do so, it collects a variety of information from the same individuals and households over a two and one-half year period--their incomes, their jobs and employment problems, their participation in government programs, their health, education, and household composition, and so on. In short, SIPP is an integrated survey with many data purposes, but one central goal: monitoring the well-being of our society.

As the Nation approaches the end of the 20th century, it is clear that our society is more complex than ever. To keep track of it we need a modern survey instrument. The SIPP is unlike any other household survey ever designed and quite capable of responding to today's and tomorrow's policy needs.

## NOTES

- 1/ For a comprehensive discussion of the CPS see John E. Bregger, "The Current Population Survey: A Historical Perspective and BLS's Role," Monthly Labor Review, June 1984.
- 2/ See Current Population Reports, Series P-60, No. 152, Characteristics of the Population Below the Poverty Level: 1984, U.S Bureau of the Census, June 1986.
- 3/ For a summary of the findings from the PSID, see Greg J. Duncan, Years of Poverty, Years of Plenty, Survey Research Center, Institute for Social Research, the University of Michigan, 1983.
- 4/ Ibid.
- 5/ See Current Population Reports, Series P-60, No. 150, Characteristics of Households and Persons Receiving Selected Noncash Benefits: 1984, U.S. Bureau of the Census, November 1985.
- 6/ See Estimates of Poverty Including the Value of Noncash Benefits: 1984, Technical Paper 55, U.S. Bureau of the Census, August 1985.
- 7/ See Martin David, "The Distribution of Income in the United States: Implications for the Design of the SIPP Panel," Journal of Economic and Social Measurement, Vol. 13, Num. 3 and 4, December 1985.
- 8/ See Joseph A. Pechman, Who Paid the Taxes, 1966-85?, Studies of Government Finance, The Brookings Institution (Washington, D.C.), pp. 4-5.
- 9/ Timothy Smeeding, "The Scientific Potential of SIPP: Its Content and Methods Regarding Fringe Benefits, Noncash Income, and the Value of Government Services," Journal of Economic and Social Measurement, Vol. 13, Num. 3 and 4, December 1985.
- 10/ See Daniel H. Weinberg, "Filling the 'Poverty Gap': Multiple Transfer Program Participation," Journal of Human Resources, Vol. 20, No. 1, Winter 1985.
- 11/ The Bureau of Labor Statistics from time to time has published reports related to this topic. See Linking Employment Problems to Economic Status, Bureau of Labor Statistics, Bulletin 2201, June 1984.
- 12/ See Kim Clark and Lawrence Summers, "Labor Market Dynamics and Unemployment," Brookings Papers on Economic Activity, No. 1, 1979.
- 13/ See Gail R. Wilensky, "SIPP and Health Care Issues," Journal of Economic and Social Measurement, Vol. 13, Num. 3 and 4, December 1985.
- 14/ See Michael R. Olneck, "Critique of Questions Pertaining to Education in SIPP," Journal of Economic and Social Measurement, Vol. 13, Num. 3 and 4 December 1985.

- 15/ See Mary Jo Bane and James Welsh, "COP's Potential Contributions to Policy Research on Children," Journal of Economic and Social Measurement, Vol. 13, Num. 3 and 4, December 1985.
- 16/ See Harold W. Watts, "The Scientific Potential of SIPP for Analysis of Living Arrangements for Families and Households," Journal of Economic and Social Measurement, Vol. 13, Num. 3 and 4, December 1985.
- 17/ See Current Population Reports, Series P-70, No. 7, Household Wealth and Asset Ownership: 1984, U.S. Bureau of the Census, July 1986.
- 18/ For background information about the ISDP see Martynas Ycas and Charles Lininger, "The Income Survey Development Program: Design Features and Initial Findings," Social Security Bulletin, Vol. 44, No. 11, November 1981.
- 19/ For a more complete description of the SIPP methodology see Dawn Nelson, David B. McMillen, and Daniel Kasprzyk, "An Overview of the Survey of Income and Program Participation: Update 1," SIPP Working Paper, No. 8401, December 1985, or see the appendixes of any Current Population Reports, Series, P-70, Economic Characteristics of Households in the United States, U.S. Bureau of the Census.
- 20/ This typology is taken from an internal Census Bureau document prepared by Roger Herriot, Senior Demographic and Housing Analyst.