

**THE SURVEY OF INCOME AND
PROGRAM PARTICIPATION**

**ESTIMATES OF THE UNINSURED
POPULATION FROM THE SURVEY
OF INCOME AND PROGRAM
PARTICIPATION: SIZE,
CHARACTERISTICS, AND THE
POSSIBILITY OF ATTRITION BIAS**

No. 161

**K. Swartz
The Urban Institute**

U. S. Department of Commerce BUREAU OF THE CENSUS

Estimates of the Uninsured Population
From the Survey of Income and Program
Participation: Size, Characteristics,
and the Possibility of Attrition Bias

by Katherine Swartz

February 1990

Funding for this paper was provided by the Aetna Casualty and Life Pension and Welfare Benefits Administration of the U.S. Department of Labor through a contract to the RAND Corporation with a subcontract to The Urban Institute. The opinions and conclusions expressed herein are solely those of the author and should not be construed as representing the opinions or policy of Aetna Casualty and Life Insurance Company, the RAND Corporation, the Urban Institute, any agency of the United States Government or the sponsors of either the RAND Corporation or the Urban Institute. The author is solely responsible for the accuracy of the statements or interpretations contained herein. The author gratefully acknowledges the comments and suggestions of Sharon Long, Timothy McBride, and Arleen Leibowitz; errors that may remain in the paper are solely those of the author. Taube Wilson provided the computer programming support.

THE URBAN INSTITUTE 2100 M Street N.W. / Washington, D.C. 20037/
(202) 833-7200

The Survey of Income and Program Participation (SIPP) interviews a sample of people every four months over a three year period. This longitudinal survey is ideal for examining behavior of people over time, particularly behavior that is thought to involve changes for some people but not everyone. Among the types of behavior that have been studied with the SIPP are participation in the labor force, the Aid to Families with Dependent Children (AFDC) cash assistance program, and Medicaid.

The SIPP has also been used to produce point-in-time estimates of the total number of people without health insurance (Census documents). However, the SIPP uninsured estimates are smaller than the more frequently used estimates of the uninsured based on the Current Population Survey (CPS). The question of why the two surveys' estimates differ has not been adequately answered. We need to understand why the aggregate estimates differ because cost estimates of various proposals to provide coverage to the uninsured depend on the aggregate number. Because the efficacy of each of the proposals to expand coverage to the uninsured depends on knowing who the uninsured are, we also need to know whether estimates of the uninsured population's characteristics differ between the two surveys.

This paper is an attempt to answer these questions. I examine the distributions of four characteristics of people without any type of health insurance at five points in time covered by the 1984 Panel (or first sample)¹ of the SIPP: September 1983, March 1984, September 1984, March 1985, and September 1985. The distributions of the age, region of residence, family income, and labor force status of the uninsured found in the SIPP are compared to the same distributions of the uninsured found in the March 1984 Current Population Survey (CPS). I also use two different subsets of the SIPP sample for estimating the uninsured population's size in order to examine the possible bias caused by the declining participation in the SIPP, a phenomenon known as attrition. The two subsets of the SIPP sample are 1) the people who have data for the first eight interviews (covering 32 months) of the survey--this subset is referred to as the 32 month SIPP file;² and 2) the people who have data for the interviews covering the five points in time examined in this paper (September 1983, March and September 1984, March and September 1985). A person does not have to be in the SIPP for all of the five months we

¹The SIPP is actually a multi-panel survey--that is, a group of people, known as a panel, are selected for the survey and are interviewed every four months for two and a half years thereafter. The first panel was interviewed for the first time in October 1983 and is known as the 1984 panel. Subsequent panels were interviewed for the first time in February of each year and are known as the 1985, 1986, and so forth panels. The separate panels permit the SIPP to be an on-going survey with concurrent administration of the survey of multiple panels. Public use data tapes for the SIPP have now been released for the 1984 and 1985 panels. The 1984 panel of the SIPP has been the data source for all of the published research on dynamic behavior that has used the SIPP, and is the data source for this paper also.

²The ninth (and last) wave of interviews was conducted on only half of the 1984 Panel sample because of budget cuts, and therefore longitudinal research with the SIPP is generally confined to the 32 month period covered for the full sample.

examine; the SIPP sample for each of the months simply consists of those people present then.³ I refer to this subset of the SIPP as the monthly SIPP file. The 32 month SIPP file has a smaller number of respondents than does the monthly SIPP file.

In the next section I discuss and compare the distributions of the characteristics of the uninsured from the two different SIPP subsamples and the CPS. In the following section I discuss why the 32 month SIPP file may suffer from what is known as attrition bias⁴ --a bias in the sample over time due to the fact that the people who leave the sample are not random with respect to some characteristics.⁵ In this case, there may be attrition bias in the estimates of the uninsured based on the 32 month SIPP file because the people who leave the SIPP are not random with respect to their health insurance coverage. After discussing why such attrition bias may exist, I also discuss why it is important to determine if such bias exists. In the last section I comment on how the findings in this paper might affect estimates of the dynamics of health insurance coverage based on the SIPP.

Comparing the Uninsured Population Estimates

The SIPP has several subsets of the sample which can be used for analyses depending on the question of interest. Studies of the dynamics of participation in the AFDC program (Ruggles, 1989) or in Medicaid (Short, Cantor, and Monheit, 1988) have used the subset of the sample that is present for 32 consecutive months, hereafter referred to as the SIPP 32 month file. Studies that have used the SIPP simply to report on the population at a point in time have used the subset of the sample that is present at the point in time being studied (see various Census documents pertaining to the SIPP). Because not everyone who is in the SIPP at one point in time is present for an earlier or later interview, the Census created monthly weights for the respondents to each interview. Thus, a respondent will have a separate weight for each month in which he or she is in the SIPP; the weights generally differ from month to month to reflect who else responded to the interview for that month.

The respondents in the SIPP 32 month file, in contrast, are by definition present for all

³This subset actually comes from what is known as the SIPP longitudinal file rather than the cross-sectional files. The reason for choosing for longitudinal file is that more work has gone into editing it and imputing answers for missing variable items. In particular, imputation procedures for people who may be missing from the sample for one interview take into account the person's answers in the previous and subsequent interviews in the longitudinal file. The cross-sectional files, in contrast, have imputed answers for missing data that rely on other, similar people's answers. Generally, the behavior of a person before and after a particular point in time is more accurate in indicating that person's behavior than is the behavior of someone who has the same demographic and socioeconomic characteristics.

⁴For further information on attrition bias, see Hausman and Wise (1979).

⁵The monthly SIPP file may also suffer from attrition bias but the bias may not be as pronounced as the bias in the 32 month SIPP file because respondents can come back into the monthly sample after an absence--something which cannot happen in the 32 month file.

eight interviews covering each of the 32 months.⁶ The only exception to this is that people who die or enter institutions leave the file. Consequently, the 32 month file contains only one weight per respondent for analyses covering the 32 month period. By summing the respondents' weights, then, we obtain the size of the U.S. population. However, the 32 month file is not intended to provide total population estimates as time passes during the 32 months because people cannot join the file to take the place of the people who die or are institutionalized.⁷ Thus, the population estimates from the 32 month SIPP file become smaller as time passes, and the respondents represent the total U.S. population only in September 1983.⁸

In this study, I have restricted the sample of interest to the non-elderly because I wanted to be able to compare the SIPP estimates of the uninsured population with estimates from the CPS, which have almost always been done for the non-elderly population. For the monthly SIPP file, restricting the sample to the non-elderly was not a problem--I simply selected only those people who were younger than 65 at that month. For the SIPP 32 month file, I eliminated anyone who was 65 or older at the beginning of the 32 month period, and then as someone became 65 years of age the person was dropped from the sample. Retaining people in the file until their 65th birthday is analogous to retaining people in the 32 month file until they died or were institutionalized.

Table 1 shows the estimates of the total non-elderly population and estimates of the total numbers of non-elderly people without health insurance at each of the five points in time examined here: September 1983, March and September 1984, and March and September 1985. The estimates are from the SIPP 32 month file, the monthly SIPP file, and the March 1984 CPS. Both SIPP subsamples' estimates of the uninsured are included because each has been the source for various studies of the uninsured. The estimate of the total population based on the 32 month file declines over the dm period we are observing because, as I explained earlier, there is only one weight for each respondent in the file. The monthly SIPP file does have separate weights for each month so the population estimates are intended to be correct for each month. Thus, the monthly SIPP file and CPS estimates of the total nonelderly population in March 1984 are identical by intention--the monthly SIPP is benchmarked to the CPS.

⁶The SIPP sample is divided into four "rotation" groups so that in any given month only a quarter of the sample is actually interviewed and questioned about the preceding four months.

⁷The 32 month file was created in order to provide data for studies of a variety of "life cycle" events that can only be studied with longitudinal data from a sample of the population.

⁸The first calendar month in which the 1984 Panel actual contains data for all four rotation groups is September 1983.

Table 1

Estimates of the Total Non-Elderly Population
and Total Number of People Without Health
Insurance from the SIPP and CPS

Estimates of Total Non-Elderly Population

<u>Month and Year</u>	<u>SIPP 32 Month File</u>	<u>Monthly SIPP File</u>	<u>CPS</u>
September 1983	205.5 m	205.1 m	
March 1984	203.2	205.6	205.6
September 1984	200.5	206.4	
March 1985	198.6	207.2	
September 1985	196.8	207.9	

Estimates of Total Number of Non-Elderly
without Health Insurance

<u>Month and Year</u>	<u>SIPP 32 Month File</u>	<u>Monthly SIPP File</u>	<u>CPS</u>
September 1983	30.3 m	32.2 m	
March 1984	28.2	29.8	35.0
September 1984	28.1	30.0	
March 1985	27.3	29.4	
September 1985	28.1	31.0	

The second point to notice about the estimates in Table 1 is that the number of people without health insurance from the SIPP 32 month file is always less than the estimate from the monthly SIPP file. The difference between the two estimates varies between 1.6 million and 2.9 million (differences of 5.7 percent to 10.3 percent). The smallest difference occurs in March 1984 and then the difference steadily increases through September 1985. Because the SIPP 32 month file is not intended to be representative of the general population after September 1983, this widening difference should not be surprising. When the SIPP estimates are compared with the March 1984 CPS estimate of the total number of uninsured, the CPS estimate is about 5 million larger than the SIPP estimate based on the subset present in March 1984 and about 7 million larger than the 32 month file. We will return to a discussion of Table 1 in the next section.

Table 2 shows the age distributions for the uninsured as estimated from the two SIPP subsets and the March 1984 CPS. The age distributions of the uninsured population from each SIPP subset are virtually identical. Thus, while the two subsets of the SIPP yield different absolute numbers of uninsured at the same points in time, they are consistent in terms of distribution of the ages of the uninsured. When the SIPP estimates of the age distributions of the uninsured are compared to the CPS based estimate, the SIPP age distribution has proportionately more children and 25 to 34 year old adults, but fewer 55 to 64 year old adults than does the CPS. There is not an obvious explanation for why these age cohorts differ between the estimates of the uninsured based on the SIPP and the CPS. In particular, given that SIPP surveys sample members every four months about their participation in various programs, we might expect the SIPP to have proportionately more children and their parents covered by Medicaid and therefore fewer such uninsured than the CPS.

Table 3 indicates the regions of residence of the uninsured populations from the two subsets of the SIPP in the five months we are examining, and the March 1984 CPS. As with the age distribution comparisons, the aggregate numbers of people without health insurance in each of the nine geographic regions are smaller when the SIPP 32 month file is used than when the monthly SIPP file is used. However, the actual distributions are similar between the two SIPP subsets. Also, the two SIPP subsets' estimates of the proportions of each region's population that is uninsured are not consistently different.

Because the March 1984 CPS estimate of the total number of people without health insurance is about 5 million larger than the March 1984 SIPP estimate, it is not surprising that the proportions of the population in each region without health insurance are higher with the CPS data than with the SIPP data. However, the distribution of the uninsured according to the CPS is greatly different from that of the SIPP. Only in the Mid-Atlantic and Mountain regions do we observe moderately large differences between the two surveys. It is not clear why these two regions should differ between the surveys.

Table 4 shows the distribution of family income relative to the poverty level for the uninsured populations from the two SIPP sample subsets and the March 1984 CPS. Table 4 continues to provide evidence that there is remarkable consistency across the two SIPP subsets and the CPS in terms of the income distribution of the uninsured.

Table 2

Age Distribution of Non-elderly Uninsured

<u>Age</u>	<u>Sept. '83</u>			<u>SIPP - Sept. '83</u>					
	<u>SIPP 32 Mo. File</u>								
<18	11.2	million	37.0%	11.8	million	36.7%			
18-24	6.8	million	22.3%	7.2	million	22.5%			
25-34	5.8	million	19.0%	6.5	million	20.2%			
35-44	2.8	million	9.3%	2.9	million	9.0%			
45-54	2.0	million	6.5%	2.0	million	6.2%			
55-64	1.8	million	5.9%	1.7	million	5.3%			
TOTAL	30.3	million	100?*	32.2	million	100%			
	<u>March '84</u>			<u>SIPP - March '84</u>			<u>CPS - March '84</u>		
	<u>SIPP 32 Mo. File</u>								
<18	10.1	million	35.8%	10.6	million	35.6%	11.6	million	33.0%
18-24	6.6	million	23.2%	6.7	million	22.4%	8.3	million	23.6%
25-34	5.3	million	18.7%	5.9	million	19.9%	6.2	million	17.7%
35-44	2.6	million	9.4%	2.8	million	9.4%	3.4	million	9.7%
45-54	2.0	million	7.0%	2.1	million	6.9%	2.7	million	7.7%
55-64	1.6	million	5.8%	1.7	million	5.7%	2.9	million	8.3%
TOTAL	28.2	million	100%	29.8	million	100%	35.0	million	100%
	<u>Sept. '84</u>			<u>SIPP - Sept. '84</u>					
	<u>SIPP 32 Mo. File</u>								
<18	10.0	million	35.5%	10.8	million	35.9%			
18-24	6.5	million	23.0%	6.7	million	22.3%			
25-34	5.3	million	18.8%	6.0	million	19.8%			
35-44	2.8	million	10.0%	3.1	million	10.3%			
45-54	1.9	million	6.8%	2.0	million	6.5%			
55-64	1.7	million	5.9%	1.6	million	5.3%			
TOTAL	28.1	million	100%	30.0	million	100%			
	<u>March '85</u>			<u>SIPP - March '85</u>					
	<u>SIPP 32 Mo. File</u>								
<18	8.9	million	32.5%	10.0	million	33.9%			
18-24	6.5	million	23.8?4	6.7	million	22.7%			
25-34	5.3	million	19.5%	5.9	million	20.1%			
35-44	2.9	million	10.6%	3.0	million	10.4%			
45-54	1.9	million	7.1%	2.1	million	7.0%			
55-64	1.7	million	6.4%	1.8	million	6.0%			
TOTAL	27.3	million	100%	29.4	million	100%			
	<u>Sept. '85</u>			<u>SIPP - Sept. '85</u>					
	<u>SIPP 32 Mo. File</u>								
<18	9.2	million	32.9%	11.0	million	35.4%			
18-24	6.4	million	22.8%	6.7	million	21.5%			
25-34	5.7	million	20.2%	6.4	million	20.5%			
35-44	3.0	million	10.9%	3.1	million	10.1%			
45-54	1.9	million	6.9%	2.1	million	6.7%			
55-64	1.8	million	6.4%	1.8	million	5.7%			
TOTAL	28.1	million	100%	31.0	million	100%			

Table 3

Region of Residence of Non-elderly Uninsured

<u>Region</u>	Sept. '83		As % of		As % of	
	<u>SIPP 32 Mo. File</u>		<u>Population</u>	<u>SIPP - Sept. '83</u>	<u>Population</u>	
Now England	1.0 million	3.2%	9.6%	0.9 million	2.9%	9.1%
Mid Atlantic	3.2 million	10.7%	9.7%	3.6 million	11.2%	10.7%
E. N. Central	4.9 million	16.3%	12.2%	4.6 million	14.9%	12.4%
W. N. Central	1.6 million	5.3%	11.9%	1.6 million	5.0%	13.5%
S. Atlantic	5.4 million	17.9%	16.7%	5.8 million	18.2%	17.4%
E. S. Central	3.0 million	9.9%	19.7%	3.1 million	9.7%	20.0%
W. S. Central	4.1 million	13.5%	21.0%	4.8 million	14.9%	23.0%
Mountain	1.5 million	5.1%	17.5%	1.5 million	4.8%	17.8%
Pacific	5.5 million	18.1%	17.2%	6.0 million	18.5%	18.6%
TOTAL	30.3 million	100%	14.8%	32.2 million	100%	15.7%

	March '84		As % of		As % of	
	<u>SIPP 32 Mo. File</u>		<u>Population</u>	<u>SIPP - March '84</u>	<u>Population</u>	
Now England	1.0 million	3.5%	9.9%	0.9 million	2.9%	8.7%
Mid Atlantic	2.9 million	10.1%	8.7%	3.2 million	10.8%	9.6%
E. N. Central	4.5 million	15.9%	11.2%	4.6 million	15.4%	11.7%
W. M. Central	1.6 million	5.5%	11.7%	1.5 million	5.1%	12.3%
S. Atlantic	5.1 million	18.2%	15.9%	5.7 million	19.0%	17.0%
E. S. Central	3.0 million	10.8%	20.0%	3.1 million	10.4%	19.7%
W. S. Central	3.9 million	13.8%	20.0%	4.4 million	14.8%	21.4%
Mountain	1.1 million	3.8%	12.5%	1.1 million	3.8%	13.5%
Pacific	5.2 million	18.5%	16.5%	5.3 million	17.7%	16.4%
TOTAL	28.2 million	100%	13.9%	29.8 million	100%	14.5%

	<u>March '84 CPS</u>		As % of
			<u>Population</u>
Now England	1.3 million	3.7%	11.9%
Mid Atlantic	4.5 million	12.9%	14.1%
E. N. Central	5.1 million	14.6%	13.9%
W. N. Central	2.2 million	6.3%	14.8%
S. Atlantic	6.3 million	18.1%	18.9%
E. S. Central	2.5 million	7.0%	18.9%
W. S. Central	5.1 million	14.5%	22.2%
Mountain	2.3 million	6.6%	21.4%
Pacific	5.7 million	16.1%	18.7%
TOTAL	35.0 million	100%	17.0%

	Sept. '84		As % of		As % of	
	<u>SIPP 32 Mo. File</u>		<u>Population</u>	<u>SIPP - Sept. '84</u>	<u>Population</u>	
Now England	1.0 million	3.5%	9.0%	0.9 million	3.2%	9.3%
Mid Atlantic	2.7 million	9.6%	8.3%	3.0 million	10.2%	9.0%
E. N. Central	4.4 million	15.7%	11.2%	4.7 million	15.6%	11.8%
W. N. Central	1.6 million	5.8%	12.2%	1.7 million	5.6%	13.3%
S. Atlantic	5.0 million	17.7%	15.7%	5.6 million	18.6%	16.5%
E. S. Central	3.0 million	10.7%	20.0%	3.0 million	9.9%	19.0%
W. S. Central	3.9 million	13.7%	20.0%	4.3 million	14.2%	20.9%
Mountain	1.1 million	4.0%	13.2%	1.2 million	3.9%	14.1%
Pacific	5.5 million	19.5%	17.7%	5.6 million	18.8%	17.7%
TOTAL	28.1 million	100%	14.0%	30.0 million	100%	14.6%

Table 3 (continued)

	March '85 <u>SIPP 32 Mo. File</u>		As % of <u>Population</u>	As t of <u>SIPP - March '85</u>		<u>Population</u>
Now England	0.9 million	3.3%	9.1%	1.0 million	3.4%	9.6%
Mid Atlantic	2.9 million	10.6%	9.0%	3.3 million	11.3%	9.7%
E. N. Central	4.2 million	15.5%	10.8%	4.4 million	14.8%	11.0%
W. N. Central	1.5 million	5.4%	11.2%	1.7 million	5.7%	12.7%
S. Atlantic	4.8 million	17.7%	15.3%	5.5 million	18.6%	16.2%
E. S. Central	2.9 million	10.7%	19.5%	2.8 million	9.7%	18.3%
W. S. Central	3.8 million	13.9%	19.8%	4.1 million	13.8%	20.0%
Mountain	1.2 million	4.3%	14.2%	1.2 million	4.2%	14.8%
Pacific	5.1 million	18.7%	16.6%	5.4 million	18.4%	17.0%
TOTAL	27.3 million	100%	13.7%	29.4 million	100%	14.2%

	Sept. '05 <u>SIPP 32 Mo. File</u>		An % of <u>Population</u>	As % of <u>SIPP - Sept. 185</u>		<u>Population</u>
New England	1.0 million	3.5%	9.9%	1.0 million	3.3%	10.2%
Mid Atlantic	2.8 million	9.9%	8.7%	3.2 million	10.3%	9.3%
E. N. Central	4.5 million	16.0%	11.7%	4.7 million	15.1%	11.9%
W. N. Central	1.6 million	5.8%	12.6%	1.8 million	5.7%	13.4%
S.. Atlantic	5.2 million	18.4%	16.5%	5.9 million	19.1%	17.5%
E. S. Central	3.0 million	10.5%	20.0%	3.0 million	9.8%	19.4%
W. S. Central	3.8 million	13.5%	19.9%	4.4 million	14.1%	20.9%
Mountain	1.2 million	4.4%	15.4%	1.5 million	4.8%	17.5%
Pacific	5.1 million	18.1%	16.8%	5.5 million	17.8%	17.4%
TOTAL	28.1 million	100%	14.3%	31.0 million	100%	14..9%

Table 4

Income Distribution of Non-elderly Uninsured

Family Income
Relative to
Poverty

	<u>Sept. '83 - 32</u>	<u>Mo. SIPP</u>	<u>Sept. '83 SIPP</u>	
<poverty	11.2 million	37.1%	11.7 million	36.3%
1.-1.49	5.5 million	18.0%	5.8 million	18.1%
1.5-1.99	3.9 million	12.9%	4.3 million	13.5%
2.-2.99	4.9 million	16.2%	5.1 million	15.8%
> 3.	4.8 million	15.9%	5.2 million	16.2%
TOTAL	30.3 million	100%	32.2 million	100%

March '84

	<u>SIPP - 32 Mo. File</u>		<u>SIPP - March '84</u>		<u>CPS March '84</u>	
<poverty	9.9 million	34.9%	10.1 million	33.9%	12.5 million	35.6%
1.-1.49	4.8 million	17.1%	5.3 million	17.8%	5.9 million	16.7%
1.5-1.99	3.8 million	13.5%	4.1 million	13.7%	4.4 million	12.6%
2.-2.99	5.3 million	18.6%	5.3 million	17.7%	5.4 million	15.5%
> 3.	4.5 million	15.8%	5.0 million	16.7%	6.9 million	19.6%
TOTAL	28.2 million	100%	29.8 million	100%	35.0 million	100%

Sept. '84 - 32 Mo. SIPP

			<u>Sept. '84 SIPP</u>	
<poverty	9.2 million	32.6%	10.0 million	33.2%
1.-1.49	5.8 million	20.6%	5.7 million	19.1%
1.5-1.99	4.1 million	14.6%	4.6 million	15.5%
2.-2.99	5.2 million	18.5%	5.4 million	17.8%
> 3.	3.9 million	13.8%	4.3 million	14.3%
TOTAL	28.1 million	100%	30.0 million	100%

March '85 - 32 Mo. SIPP

			<u>March '85 SIPP</u>	
<poverty	8.7 million	32.1%	9.4 million	32.2%
1.-1.49	5.2 million	18.9%	5.4 million	18.4%
1.5-1.99	4.5 million	16.5%	4.8 million	16.2%
2.-2.99	4.5 million	16.5%	5.0 million	17.0%
> 3.	4.4 million	16.1%	4.7 million	16.1%
TOTAL	27.3 million	100%	29.4 million	100%

Sept. '85 - 32 Mo. SIPP

			<u>Sept. '85 SIPP</u>	
<poverty	9.0 million	32.0%	10.3 million	33.2%
1.-1.49	5.3 million	18.8%	5.7 million	18.5%
1.5-1.99	4.4 million	15.6%	4.6 million	14.9%
2.-2.99	4.9 million	17.3%	5.2 million	16.8%
> 3.	4.6 million	16.4%	5.1 million	16.6%
TOTAL	28.1 million	100%	31.0 million	100%

Table 5

Employment Status of Uninsured 18-64 Year Old Adults

<u>Status</u>	<u>Sept. '83 - 32 Mo. SIPP</u>		<u>Sept. '83 - SIPP</u>			
employed	10.3 million	53.9%	10.8 million	52.9%		
unemployed	3.7 million	19.7%	4.1 million	20.2%		
out of labor force	5.1 million	26.6%	5.5 million	26.8%		
Total	19.1 million	100%	20.4 million	100%		
	<u>March '84 - 32 Mo. SIPP</u>		<u>March '84 - SIPP</u>		<u>March '84 - CPS</u>	
employed	9.8 million	53.8%	10.2 million	53.0%	13.2 million	56.5%
unemployed	3.2 million	17.6%	3.5 million	18.2%	2.8 million	12.1%
out of labor force	5.2 million	28.6%	5.5 million	28.8%	7.3 million	31.4%
Total	18.1 million	100%	19.2 million	100%	23.4 million	100%
	<u>Sept. '84 - 32 Mo. SIPP</u>		<u>Sept. '84 - SIPP</u>			
employed	10.3 million	56.7%	10.8 million	56.1%		
unemployed	2.9 million	15.8%	3.0 million	15.8%		
out of labor force	5.0 million	27.5%	5.4 million	28.2%		
Total	18.1 million	100%	19.3 million	100%		
	<u>March '85 - 32 Mo. SIPP</u>		<u>March '85 - SIPP</u>			
employed	10.6 million	57.5%	10.8 million	55.6%		
unemployed	3.0 million	16.3%	3.3 million	17.1%		
out of labor force	4.8 million	26.3%	5.3 million	27.2%		
Total	18.4 million	100%	19.4 million	100%		
	<u>Sept. '85 - 32 Mo. SIPP</u>		<u>Sept. '85 - SIPP</u>			
employed	11.0 million	58.2%	11.5 million	57.6%		
unemployed	2.9 million	15.1%	3.0 million	15.1%		
out of labor force	5.0 million	26.7%	5.5 million	27.3%		
Total	18.9 million	100%	20.0 million	100%		

Finally, Table 5 shows the employment status of the 18 to 64 year-old adults without health insurance from the two SIPP sample subsets and the March 1984 CPS. Again, the distributions are remarkably consistent across the data sets. The rising number of uninsured adults who are employed and the simultaneous decline in the number who are unemployed between September 1983 and September 1985 reflects the improving economy during that time period. Interestingly, the CPS labor force distribution of uninsured adults shows a larger proportion out of the labor force and a larger proportion employed than does the SIPP. Some of this difference must be due to differences in the way the two surveys categorize responses to questions about work activities during the month.

In sum, the distributions of age, region of residence, family income relative to poverty, and employment status of the uninsured are quite similar across the two SIPP sample subsets and the March 1984 CPS. With the exception of the first quarter of the SIPP (September 1983), the SIPP estimates of the uninsured based on the monthly SIPP file are about 5 million less than the March 1984 CPS estimate. If the SIPP 32 month file subset of the sample is used, the SIPP estimate of the uninsured in September 1983 is lower than the March 1984 CPS estimate by yet another 2 million.

Health Insurance and Attrition Bias

In order to obtain correct estimates of the distributions of the population's age, sex, race, and geographic residence, most national surveys create weights for the survey respondents based on their age, sex, race, and location of residence. These simple characteristics are used to calculate weights because they are information collected by the Census every ten years for the entire population and because they are immune to external forces other than extreme factors such as war.

Longitudinal sampling--i.e., a survey that follows the same sample over a period of time rather than just a one time survey--poses special problems in the calculation of weights for a sample's respondents because of the loss of respondents over time. The loss of sample respondents is known as attrition. Respondents leave a sample over time because of death, or address changes that are not provided to the survey, or because the respondent no longer wishes to participate in the survey. As long as there is a random pattern to the types of people who leave a longitudinal sample, the attrition from the sample does not bias the subsequent estimates of the size or characteristics of the population. The SIPP, as well as other large national longitudinal surveys, corrects for the loss of respondents by changing the weights of the remaining respondents along the characteristics of age, sex, race, and geographic residence. The Bureau of the Census is very concerned with correcting for sample attrition.⁹ However, it cannot correct for attrition linked to a particular characteristic when we do not know what the

⁹McArthur and Short (1985) showed that attrition rates from the SIPP are higher for young adults, male, blacks, Hispanics, never married persons, primary individuals and persons not related to the reference person in the household.

true underlying distribution of the population is with respect to that characteristic. This is particularly troublesome when the characteristic is one in which we are very interested in observing changes, such as the income distribution of the population.

Problems from attrition arise when the people who leave a sample are not randomly leaving with respect to the behavior in which we are particularly interested. To take an extreme example, if all the people who left a longitudinal sample were people who were unemployed, we would not be able to determine what types of factors are associated with losing a job and subsequently attaining employment. In the case of the SIPP, the sample lost respondents during the two and a half years of the 1984 panel for two quite distinct reasons. First, in an effort to save money, the sample for the 1984 panel was deliberately reduced by 18 percent in March 1985--a year and a half after the panel was initiated. The sample reduction was done randomly so it would not add to attrition bias. But among the remaining sample (i.e., the sample after the 18 percent reduction), close to 20 percent of the people who were interviewed in the first round of the survey were not in the sample 32 months later (McArthur, 1988; Kasprzyk and McMillen, 1987). More than half of these respondents left within the first 12 months. It is this second group of people, who voluntarily left the SIPP, that cause concern because their attrition may be correlated with changes in job or marital status and health insurance coverage.

The impact of attrition from the SIPP is of interest because the SIPP yields a lower estimate of the uninsured population than do either the CPS or the 1987 National Medical Expenditure Survey (NMES). Some of the difference between the SIPP estimates of the size of the uninsured population and the CPS and NMES estimates can be explained by wording differences in questions about the same type of health insurance. For example, the CPS asks about Medicaid while the SIPP asks about Medicaid or any other government program that pays for medical care. The SIPP question is bound to pick up participation in state-run health care programs that pay for special services (such as mammographies) but are not equivalent to Medicaid.

An alternative potential explanation for a portion of the differences in the estimates of the number of people without health insurance is that the attrition from the SIPP is not random. In particular, if the people who leave the sample are also more likely to be uninsured, or to become uninsured at the same time they leave the sample, then the SIPP estimates of the uninsured population will be biased downwards for time periods after the first interview. This issue is also important because the research being done on the dynamics of health insurance coverage with the SIPP has generally been conducted with the subset of the SIPP sample that remains in the sample for the 32 months covered by the survey. If the people who are in the SIPP sample for 32 consecutive months are representative of people who are "stayers" rather than "movers", and if being a "stayer" is correlated with not changing health insurance status as often as a "mover" does, then the 32 month SIPP file will yield biased estimates of the distribution of spells of being uninsured as well as biased estimates of the parameters of

characteristics associated with changes in health insurance status.¹⁰

A simple test of the possibility that the people who left the SIPP were more likely to be uninsured than the stayers can be conducted by comparing the estimates of the uninsured from the SIPP monthly file and the 32 month file in September 1983 (Table 1). Recall that September 1983 is the only time when the 32 month file is representative of the U.S. population. The difference between the two SIPP files' estimates of the number of uninsured, 1.9 million people, indicates that attrition bias with respect to health insurance coverage does exist in the SIPP.¹¹

Another test of whether attrition bias is present, suggested by Sharon Long, is to compare the percentages of people in the SIPP monthly file and 32 month file who remain in their initial insurance state in the months they can be observed. Table 6 presents data for the two SIPP files indicating the percent of people who were either never uninsured or were uninsured 95- 100 percent of the time they can be observed in the SIPP. The proportions of people who are insured in all the months they can be observed do not differ greatly between the two files. But the proportions who are uninsured 95 - 100 percent of the months they can be observed are significantly different between the two files. The fact that the 32 month file has a significantly smaller proportion of respondents who are almost always uninsured compared to the monthly file provides further evidence that leaving the SIPP is correlated with being uninsured, and attrition bias with respect to health insurance exits in the SIPP. Implications for Investigations of Health Insurance Dynamics

Thus, we have a "good news, bad news" situation in terms of estimating the characteristics of the population without health insurance. The "good news" is that at least two subsets of the SIPP sample provide estimates of the distributions of four characteristics of the uninsured population that are consistent with estimates of the distributions of the same characteristics from the CPS and earlier surveys that collected data about health insurance coverage (e.g., NMCES). Thus, in terms of assessing the characteristics of the uninsured at various points in time, several surveys are now providing robust estimates of the distributions of characteristics of the uninsured population.

¹⁰Clark and Speare (1988) and Speare and Avery (1989) have made similar points with regard to estimates of geographical mobility and household change using only respondents who have complete interviews in all relevant waves of the SIPP.

¹¹My thanks to Tim McBride for making me focus on the September 1983 estimates.

Table 6

**Percent of Respondents Who Were Never Uninsured and
Percent of Respondents Who Were Almost Always
Uninsured During Time Observed by SIPP**

<u>SIPP File</u>	Percent of Time Uninsured:	
	<u>None</u>	<u>95-100%</u>
32 Month File	72.2%	3.7%
Monthly File	70.5%	6.8%

The "bad news" is that there is evidence of attrition bias with respect to health insurance coverage in the SIPP--i.e., the people who leave the SIPP sample are more likely to be without health insurance than are the people who stay in the sample. One implication of this is that the 32 month file should be used with the full knowledge that the uninsured stayers in the sample may not represent the uninsured in terms of the lengths of spells of being uninsured. The 32 month file may contain a disproportionate number of uninsured people with longer spells of being without health insurance than truly exists in the population. Or it may contain a disproportionate number of people with short spells of being without health insurance. We cannot determine yet which of these scenarios is more accurate. A second implication of the attrition bias is that the 32 month file may yield biased estimates of the relative impact of factors associated with the beginning or end of observed spells of no insurance.

One approach to use in estimating the duration of spells of being without health insurance with the SIPP would be to use what is known as a "competing risk" duration model. The competing risks are that the person becomes insured or that the person leaves the sample. Thus, a spell of being uninsured can end either with the person becoming insured or leaving the sample. This approach enables us to use information from more people than just those in the 32 month file, and therefore we are not discarding information from the people who leave the SIPP sample.

REFERENCES

Clark, R. and A. Speare, Jr.: "Measuring Geographical Mobility Using Panel Data from the SIPP." Paper given at the annual meetings of the American Statistical Association, New Orleans, August 1988.

Hausman, J. and D.A. Wise: "Attrition Bias in Experimental and Panel Data: The Gary Income Maintenance Experiment," Econometrica, Vol. 47., No. 2, March 1979, pp. 455 - 473.

Kasprzyk, D. and D.B. McMillen: "SIPP: Characteristics of the 1984 Panel." Paper presented at the 1987 annual meeting of the American Statistical Association, August 1987.

McArthur, E.: "Measurement of Attrition Through the Completed SIPP 1984 Panel: Preliminary Results," Memorandum, U.S. Bureau of the Census, March 1988.

Monheit, A.C. and C.L. Schur: "The Dynamics of Health Insurance Loss: A Table of Two Cohorts", Inquiry, Vol. 25, (Fall 1988), pp. 31- 327.

Ruggles, P.: "Welfare Dependency and Its Causes: Determinants of the Duration of Welfare Spells." Paper presented at the annual meetings of the Allied Social Science Association, New York, 1988; Urban Institute Working Paper, (revised), 1989.

Short, P.F., J.C. Cantor, and A.C. Monheit: "The Dynamics of Medicaid Enrollment", Inquiry, Vol 25, (Winter 1988), pp. 504 - 516.

Speare, A., (Jr.) And R. Avery: "Measuring Household Changes at the Individual Level Using Data from SIPP," U.S. Bureau of the Census Survey of Income and Program Participation Working Paper No. 8917, July 1989.