Survey of Income and Program Participation

Welfare Dependency and Its Causes:
Determinants of the Duration of Welfare Spells

8908 85

Patricia Ruggles
The Urban Institute

May 1989

This paper was presented at the Annual Meeting of The American Economic Association, December 1988, New York, New York. The views expressed are attributable to the author and do not necessarily reflect those of the Census Bureau.

TABLE OF CONTENTS

Modeling the Duration of Welfare Spells..... 3

Data and Methodology..... 4

Estimates of the Duration of Welfare Spells..... 6

Table 1: Life Table Survival Estimates for AFDC Spells, by Various Characteristics of Recipients..... 7

Table 2: Duration Models for Welfare Spells..... 10

References..... 14

Welfare programs serve an important role in providing both short and long term assistance to low-income families and individuals. The Aid to Families with Dependent Children (AFDC) program, for example, the basic cash assistance program for low-income families with children, aids about 11 million people each month. The impacts of AFDC recipiency on work incentives and family composition have been studied in depth by a broad range of policy analysts.[1]

More recently, however, concern has also focused on the dynamics of welfare recipiency. Much of this concern has arisen in the wake of an influential study by Mary Jo Bane and David Ellwood (Bane and Ellwood (1983)), which found that although most new AFDC cases received welfare for only a relatively short period of time (two years or less) a minority remained on welfare over a much longer period. These long-spell cases were found to account for a sizable proportion of the total caseload at any given point in time, and to occasion the bulk of welfare costs. These findings have led to an increased interest on the part of analysts in modeling the determinants of welfare spell durations.

Studies of the dynamics of welfare recipiency are a fairly recent phenomenon, however, at least partly because detailed data on spell lengths and personal characteristics of recipients have been hard to find. Studies of AFDC participation by Hutchens (1981) and Plotnick (1983) examined transitions into and out of AFDC, but did not consider issues relating to spell length directly. The first analyses to investigate spell durations explicitly were the Bane and Ellwood study mentioned above, which used 12 years of data on AFDC participation from the Panel Study of Income Dynamics (PSID), and a study by June O'Neill, Douglas Wolf, Laurie Bassi, and Michael Hannan (1984), which used not only the PSID but also data from the National Longitudinal Survey (NLS) and from administrative records on AFDC case openings and closings. Both of these studies focused primarily on the determinants of spell durations, investigating the specific impacts of demographic, economic, and program-related variables. More recently, Ellwood (1986) has updated the PSID results, while O'Neill, Bassi and Wolf (1985) have further examined AFDC spells observed in the NLS, using

---

1. See for example Goodwin and Mahard (1986), or Ellwood and Bane (1984) for discussion of this literature.

several variations on discrete duration dependence models to test for increased probability of continued participation as spells lengthen. Finally, a recent paper by Rebecca Blank has introduced a more rigorous definition of welfare dependency--essentially, a decrease in the conditional probability of leaving welfare as spell duration increases--and has examined dependency using monthly data on participation from the control group for the Seattle/Denver Income Maintanance Experiments (SIME/DIME). Using this definition, she finds little evidence of welfare dependence across a variety of model specifications.[2]

This paper examines the dynamics of welfare receipt and the determinants of welfare spell durations using newly available panel data from the Survey of Income and Program Participation (SIPP). The paper considers the dynamics of welfare recipiency in general, and, unlike for example Blank's study, does not examine a formal model of dependence defined as a change in the conditional probability of a welfare exit. We hope to extend the work in this direction at some future point, but the focus of this more preliminary examination is on the characteristics of recipients as they influence spell durations.

The SIPP data used in this analysis provide detailed monthly information on the demographic and economic characteristics of families and households on a month by month basis. With the exception of Blank's SIME/DIME data, which are both rather old and limited to a very non-representative set of sites, all of the other dynamic participation models seen in the literature are based on annual data. In a monthly program like AFDC use of annual data can bias estimated spell durations significantly. In addition, it is more difficult to observe the specific characteristics of the AFDC unit and household at the time of spell entry or exit using annual data, particularly where changes in these variables occur during the year. The SIPP thus represents an opportunity for substantial improvements in our estimates of AFDC spell durations.

---

2. Two other recent papers consider issues relating to welfare dynamics without modeling dependency explicitly: John Fitzgerald (1988) has used data from the Survey of Income and Program Participation (SIPP) to examine the impacts of marriage opportunities on AFDC exit rates, and Roberton Williams and Patricia Ruggles have also used SIPP data to examine welfare transitions more generally.

Modeling the Duration of Welfare Spells

As discussed above, a fairly large number of authors have modeled aspects of welfare program participation over the past several years. Such models typically see the decision to participate (or to continue participating) in a welfare program as an issue of choice: a woman (or couple) chooses to participate if the utility of doing so exceeds the utility derived from not doing so—i.e.,

$$U_w > U_n.$$

The participation function, then, may be written simply as

$$\Phi = U_w - U_n > 0.$$

A dynamic component may be added to this model simply by assuming that the participation decision and its components, U_w and U_n, are reassessed in each period, so that

$$\Phi_t = U_{wt} - U_{nt} > 0.$$

If utility is a function of income and leisure, as is generally assumed, plus some specific household characteristics that determine the shape of the underlying function, then a generalized utility function may be written

$$U = U(H,Y,X)$$

where H = hours worked (negative leisure), Y = income, and X is a vector of specific household characteristics. For both U_w and U_n, however, the determinants of Y may shift considerably with changes in the X vector. For example, consider Y_n, which is a function of

$$Y_n = f_n(E_n,E_s,C,O,I),$$

where E_n = the household head's earnings, E_s = the spouse's earnings (if any), C = the child care costs necessary for the head (or if present, spouse) to work, O = other income (for example, alimony or child support), and I = the information and search costs involved in obtaining a job in the first place if either the head or the spouse does not currently have one. Similarly, Y_w will also be a function of both E_w and some other important factors:

$$Y_w = f_w(E_w,C,O,B,E), \text{ where}$$

E_w = earnings during welfare recipiency periods, for any earners in the household, C and O are as above, and B and E relate to the available welfare

programs: B = the benefit for the family's size in its state of residence, and E is a vector of family characteristics related to the state's welfare program eligibility rules.

As may be seen, both Y_n and Y_w are dependent on specific variables related to the X vector of family characteristics. These include for example the number and ages of children in the household (the primary determinants of child care costs); the presence of a spouse; the head's marital history (which is likely to affect other income such as alimony or child support); and the head's education and /or job skills (which will affect not only potential earnings but also the information costs of finding a job.)

The X vector of family characteristics may also influence the shape and/or location of the utility function more directly, if perceptions about the social acceptability of welfare program participation also affect the relative utility of welfare receipt. For some individuals, welfare recipiency may be perceived as a source of social stigma, decreasing the likelihood that, all else held constant, they will choose to participate in wefare programs. Others may be less affected—for example, some may belong to a subculture that does not regard welfare recipiency as particularly deviant, while others may simply care less about deviation from social norms in general. While it is difficult to test directly for these factors, the presence of other behaviors that deviate from social norms—for example, a birth while unmarried—may indicate a higher tolerance for stigma effects.

In essence, then, this model predicts that factors that reduce potential income from non-welfare sources such as earnings will increase spell durations, all else held constant. Additionally, to the extent that certain individuals experience less stigma as a result of welfare recipiency, they would also be expected to have longer spells.

Data and Methodology

The data used in this study are drawn from the 1984 panel of the SIPP, which follows an initial sample of about 53,000 people over a period of 32 months starting in the fall of 1983. The single biggest advantage of the SIPP is that

it collects monthly data on income, household composition, and program participation for a fairly large, representative sample of households. Because these data are longitudinal, however, month to month inconsistencies in reporting that could not be observed in a cross-sectional file become very apparent. Also, as a new file, the SIPP has not undergone the careful editing procedures that are applied to other Census Bureau data products, and particularly for longitudinal analysis, some further editing is typically necessary. The AFDC file used in this analysis, which contains 491 cases with observed AFDC spell entries, was constructed from a version of the 1984 panel file that had been substantially edited for consistency. The edits applied are described in detail in Coder and Ruggles (1988), and will not be further discussed here.

The methods used to examine the determinants of welfare spell durations in this paper apply to a dynamic version of the basic choice model discussed above. First, a survival function for welfare participation is estimated by defining $F^*(t, X_t)$ as the cumulative distribution of time on welfare, with X_t defined as a vector of relevant household characteristics and program parameters, as above, and with F^* representing the results of a series of participation decisions, Φ_1 through Φ_t. At any time t, then, $F^*(t,X_t)$ may be seen as representing the probability that the duration of welfare for someone with the given X vector of characteristics is < t. The density function associated with this distribution of survival times may be denoted $f(t, X_t)$. The survival function for participation is then simply the proportion still on welfare at time t—that is, $S(t,X_t) = 1 - F^*(t,X_t)$. The instantaneous rate of exit from welfare, or the hazard rate for exits, conditional on participation up to time T=t, is then given by

$$\lambda(t,X_t) = \lim_{\delta t \to 0} \frac{\text{prob}(t<T<t+\delta t \mid T>t, X_t)}{\delta t}$$

$$= f(t, X_t)/S(t,X_t)$$

$$= \delta(-\ln(S(t,X_t)))/\delta t.$$

If this is integrated, the survival function becomes

$$S(t,X_t) = \exp(-\int_0^t \lambda(u,X_u)du).$$

The specific functional forms of the hazard model that are estimated here include both a Weibull and a loglogistic distribution for the hazard function. The Weibull distribution is relatively easy to estimate and is therefore often chosen for survival analyses of this type, and is shown here to offer a benchmark for comparison with other studies. The loglogistic distribution was chosen because Blank, who investigated a number of possible functional forms, found that the loglogistic provided the best fit for her AFDC spell data, which appear to be distributed quite similarly to the SIPP data. In fact, in preliminary goodness-of-fit tests across the Weibull, loglogistic, exponential, and log normal distributions for our spell data the loglogistic function also appeared to provide the best fit for the SIPP results.[3]

Estimates of the Duration of Welfare Spells

Before turning to the results of the model of the determinants of duration described above, it may be instructive to examine some simpler estimates of spell durations by recipient characteristics.[4] These data, shown in Table 1, make it clear that the use of monthly data on participation does result in substantially shorter estimated spell durations than those found by Bane and Ellwood using the PSID. As the first column of Table 1 shows, more than half of all AFDC recipients have left the program by the end of the first year—the median spell length is about 11 months, in contrast to the median of about 2

---

3. See Allison (1982) and Tuma and Hannan (1984) for more discussion of modeling a time-related dependent variable in a survival function context. Blank (1986) also discusses the implications of using alternative hazard distributions.

4. The estimates presented here (and throughout the paper) are for first observed welfare spells only (although in a small number of cases very short intervals between spells were edited out, using the procedures described in Coder and Ruggles (1988)). Further examination of multiple spells will be undertaken, but the SIPP observation period is so short relative to the median spell length that in practice only a few returns to welfare can actually be observed.

Table 1

Life Table Survival Estimates for AFDC Spells, by Various Characteristics of Recipients

	All Recipients	Marital Status		Employment Status		Race		Age of Youngest Child	
		Never Married	Ever Married	Previous Employment	No Employment	Non-White	White	5 or less	6 or more
Total Number of Cases	491	160	331	143	348	183	308	327	164
Percentage of Cases Remaining (Survival Rate) After:									
4 months	76.1	84.7	71.9	65.2	80.5	81.6	72.9	77.3	73.8
	(2.0)	(2.8)	(2.6)	(4.1)	(1.9)	(3.0)	(2.6)	(2.4)	(3.5)
8 months	56.4	67.9	50.7	41.7	62.3	67.6	50.0	57.9	53.3
	(2.4)	(3.9)	(3.0)	(4.4)	(2.8)	(3.7)	(3.0)	(2.9)	(4.2)
12 months	46.5	61.2	39.5	33.6	51.7	57.0	40.1	47.1	45.2
	(2.5)	(4.3)	(3.0)	(4.4)	(3.0)	(4.2)	(3.1)	(3.1)	(4.3)
16 months	38.1	54.2	30.4	26.5	42.7	49.5	31.8	38.1	38.0
	(2.6)	(4.7)	(3.1)	(4.5)	(3.2)	(4.5)	(3.2)	(3.3)	(4.5)
20 months	32.8	45.7	26.6	22.1	37.0	41.2	28.0	31.9	34.7
	(2.7)	(5.1)	(3.1)	(4.7)	(3.3)	(4.9)	(3.2)	(3.3)	(4.7)
24 months	28.5	40.7	22.6	14.8	33.7	36.0	24.4	28.6	28.6
	(2.9)	(5.3)	(3.3)	(4.7)	(3.4)	(5.1)	(3.3)	(3.4)	(5.0)
28 months	26.3	40.7	19.3	11.5	32.0	33.3	22.3	26.7	25.6
	(3.0)	(5.3)	(3.5)	(4.7)	(3.6)	(5.4)	(3.6)	(3.6)	(5.3)

Source: Calculated from a 32 month panel of the 1984 Survey of Income and Program Participation. Survival standard errors are shown in parentheses under each estimate.

But what does this say about the 2yr period for SIPP Sample?

years found by Bane and Ellwood. As discussed earlier, the major reason for this difference is probably the fact that spells are measured in months rather than years in the SIPP data.[5]

Table 1 also demonstrates that there are indeed substantial differences in predicted spell durations for different subgroups within the population. Mothers who have never been married are likely to have considerably longer spells than the ever-married group (who are predominantly divorced or separated). The median spell duration for never-married welfare recipients is between 17 and 18 months, compared to just over 8 months for the ever-married group. Additionally, 40 percent of the never-married mothers are still receiving welfare after two years.

A second variable that appears to have a considerable impact on spell durations is employment status. Recipients who were employed in either the month before or the month of the start of the welfare spell are likely to experience much shorter spells than are those who were not recently employed.[6] The median spell length for those employed immediately before the start of the welfare spell is less than 5 months, compared to over 12 months for those who were not employed. This employment variable, which is a very basic indicator of labor-force attachment, may be picking up both a measure of commitment to work and some indication of employment-related skills.[7] Those with recent labor

---

5. This estimate is consistent with the median of about 10 months estimated by O'Neill et al. on the basis of administrative data on AFDC case openings and closings. It is lower than the median of about 18 months estimated by Blank using SIME/DIME data, but these data were not nationally representative.

6. This employment variable was constructed to parallel the "job loss" variable used in our previous work on transitions onto and off of welfare programs (see Williams and Ruggles (1987)). In that paper, we found that loss of a job was a fairly strong predictor of welfare entries in the same or the succeeding month, but we hypothesized that such entries were likely to lead to relatively short spells. Our research here confirms that hypothesis. In fact, subsidiary analyses for this project found that duration estimates were not terribly sensitive to the exact specification of the employment variable--anyone reporting employment either immediately before or during the welfare spell was likey to have a much shorter than average spell duration.

7. Ideally, a broader measure of job skills, such as education, should also be examined, but data on educational attainment are not available on the specific SIPP extract used for this study. They are available on the larger

(Footnote 7 Continued on Next Page

force experience must have at least some ability to find and hold a job, and are likely, on average, to have more such skills than those with no recent job.

Other variables examined in Table 1 include race and the age of the family's youngest child. Race does appear to make a difference, with non-whites experiencing a median spell of just under 16 months, compared to about 8 months for whites. As with the other two variables, differences in spell durations for the two subpopulations were significant at the one percent level using either a log rank or Wilcoxon rank test. Presence of a young child in the household, however, does not produce significant differences in spell durations, even though it might be expected to increase child care costs, holding down the probability of spell exits through employment.[8]

Although the results presented in Table 1 make a strong case for differences in expected durations for those in different subpopulations, they do not give any indication of the relative importance of specific variables in predicting spell durations. Table 2, which shows the outcome of the two forms of the duration model outlined above, allows us to consider the impacts of these variables on spell durations while also taking the effects of other factors into account. In addition to the four variables shown in Table 1, the duration model includes information on family type, the number of children in the AFDC unit, the age of the mother in the unit, the maximum AFDC benefit (normalized for a family of three) available in the unit's state, the unemployment rate in the unit's state, and the unit's other income. Family type is included on the theory that units that are embedded in larger households (i.e. subfamilies) may be able to draw some support, both financial and in terms of child care, from that household, increasing their liklihood of exit from AFDC. Additional children, on the other hand, directly increase AFDC benefits (which rise with family size) while indirectly decreasing the returns to work, through their

---

(Footnote 7 Continued from Previous Page)
SIPP file, and will be added to this analysis as soon as we can add them to our dataset.

8. Age cutoffs below 6 years were also examined, and were also found to produce only insignificant differences between those with and without young children.

Table 2

Duration Models for Welfare Spells

	Mean Value of Variable	Weibull Hazard	Loglogistic Hazard
Constant		2.287**	1.766**
		(0.410)	(0.416)
Race0.63	-0.215	-0.317*	
1=White		(0.145)	(0.147)
Marital Status	0.67	-0.650**	-0.527**
1=Ever married		(0.179)	(0.181)
Employment Status	0.71	0.601**	0.670**
1=no recent job		(0.135)	(0.145)
Presence of Child Under 6	0.67	-0.002	-0.003
1=yes		(0.146)	(0.152)
Family Type	0.72	0.242	0.217
1=no subfamily		(0.160)	(0.167)
Teen-aged Mother	0.75	-0.189	-0.331*
1=no		(0.183)	(0.194)
Number of Children	1.86	0.149**	0.155**
		(0.058)	(0.58)
Maximum AFDC Benefit	365.64	0.00002	0.00001
		(0.00005)	(0.00005)
Other Income	45.10	-0.0005	-0.0007
		(0.0005)	(0.0005)
Unemployment Rate	7.96	0.039	0.047
		(0.039)	(0.039)

Source: Calculated from a 32 month panel of the 1984 SIPP.

Standard errors in parentheses.

** Significant at the 1 percent level.
* Significant at the 10 percent level.

potential impacts on child care costs. Both of these effects would be expected to increase spell lengths. Teenaged mothers may have fewer job skills than older mothers, decreasing the probability that they will exit through employment and increasing the expected duration of their welfare spells. Higher AFDC benefits would be expected to increase the probability of remaining on AFDC, while the presence of other income (in this case, child support and alimony, earnings of family members other than the head, and property income) might reduce spell lengths, both by reducing the relative returns to AFDC and by increasing the family's options for support. Finally, high local unemployment rates imply a relatively difficult market for job seekers, which would also result in longer expected AFDC spells.

As Table 2 shows, both marital status and recent employment experience have a major impact on estimated spell durations even when these other factors are accounted for. Both variables are significant at the 1 percent level in both versions of the duration model, and both have large associated estimates.[9] As seen earlier, ever-married status has a negative impact on the duration of AFDC spells, while having no employment experience immediately before the spell entry increases durations. The number of children in the AFDC unit is also highly significant in both models, with larger numbers of children increasing expected spell durations, as expected.

The presence of young children has no apparent impact on durations in either model. Race is significant at the 10 percent level in the loglogistic version of the model, but not if a Weibull distribution is used. In both cases the sign is in the expected direction. The correlation matrices for these models indicate that race has less impact here than in the life tables because it is fairly highly correlated with the marital status variable (and to a lesser extent, with presence of a young child and presence of a teenaged mother). The presence of a subfamily is also not statistically significant, but has a sign as predicted above. Presence of a teen mother does appear to increase expected

9. It should be noted that the estimates shown in Table 2 refer to the likelihood of remaining on AFDC, rather than the likelihood of exiting (which is more commonly shown) and as a result all of the signs of the coefficients are the reverse of those seen in exit models.

spell lengths, but is significant only at the 10 percent level, and only in the loglogistic version of the model. Finally, maximum benefit levels, other income, and unemployment rates are all insignificant, although all have signs in the expected directions.

The results seen in Tables 1 and 2 have some important implications for an understanding of welfare dynamics. Specifically, they imply that never-married mothers and those with no job before the spell entry are likely to experience much longer welfare spells than those who have been married or who have been recently employed. As discussed above, the marital status impact may well arise, at least in part, out of different perceptions about welfare use among unmarried mothers and those in the larger society as a whole. Unmarried mothers may be more likely to belong to a subculture where welfare use is considered relatively normal—or alternatively, women who become mothers while unmarried may simply have a higher tolerance for deviations from social norms. The impacts of the employment status variable may be partly economic in nature--presumably, women who are more employable and who have higher potential wages are more likely to have recent employment experience, all else held constant. Attachment to the labor force may also indicate either a relative taste for work or a greater acceptance of social norms calling for self-support, however.

In sum, although these findings are preliminary and substantial work remains to be done, they emphasize the importance of the mother's basic socio-economic characteristics in predicting welfare spell durations. These characteristics are undoubtedly related to real differences in womens' job opportunities and potential non-welfare incomes, but they may be at least as important for their influence on perceptions of the acceptability of welfare use and the availability of other options. The fact that variables such as the age of youngest child, which are clearly related to potential net earnings, have no impact on spell lengths may indicate that potential earnings relative to benefits are not the only factor women use in deciding whether to continue participating in AFDC.

These findings also demonstrate, however, that for many mothers, particularly divorced and separated women with some recent work experience, AFDC

spells are likely to be quite short. For these women the program clearly does serve primarily as a source of very short term emergency support. Welfare recipiency is likely to last much longer for the minority of women who enter as never-married mothers, but even for this group the median spell duration is between 17 and 18 months (although about 80 percent of such mothers whose spells exceed the median are still recipients a year later). At a minimum, then, these findings suggest that attempts to decrease dependency will have the highest returns if focused on unmarried mothers and those without recent work experience, and that other recipients are likely to leave the program fairly quickly even without intervention.

References

Allison, P.D. "Discrete-Time Methods for the Analysis of Event Histories," in Sociological Methodology 1982, S. Leinhardt (ed.), San Francisco: Jossey-Bass, 1982.

Bane, Mary Jo and David T. Ellwood. "The Dynamics of Dependence: The Routes to Self-Sufficiency." Report prepared for the U.S. Department of Health and Human Services by Urban Systems Research and Engineering, Inc. Harvard University.

Coder, John F. and Patricia Ruggles. "Welfare Recipiency as Observed in the SIPP." SIPP Working Paper No. 8818. U.S. Bureau of the Census, July 1988.

Ellwood, David T. "Targeting "Would Be" Long-Term Recipients of AFDC." Report prepared for U.S. Department of Health and Human Services by Mathematica Policy Research, Inc. January, 1986.

Ellwood, David T. and Mary Jo Bane. "The Impact of AFDC on Family Structure and Living Arrangements." Report prepared for the U.S. Department of Health and Human Services. Harvard University. March 1984.

Fitzgerald, John. "The Effects of the Marriage Market and AFDC Benefits on Exit Rates from AFDC." Paper presented at the Population Association of America Meetings, New Orleans, March 1988.

Garfinkel, Irwin and Sarah McLanahan. Single Mothers and Their Children: A New American Dilemma. Washington, DC: The Urban Institute Press, 1986.

Hutchens, Robert. "Entry and Exit Transitions in a Government Transfer Program: The Case of Aid to Families with Dependent Children," Journal of Human Resources, 1981, pp. 217-237.

O'Neill, June A., Douglas Wolf, Laurie J. Bassi and Michael T. Hannan. "An Analysis of Time on Welfare." Report prepared for U.S. Department of Health and Human Services. The Urban Institute, June 1984.

O'Neill, June A., Laurie J. Bassi and Douglas A. Wolf. "The Duration of Welfare Spells," The Urban Institute, December 1985.

Plotnick, Robert. "Turnover in the AFDC Population: An Event History Analysis," Journal of Human Resources, 1983, pp. 65-81.

Tuma, N.B. and Hannan, M.T. Social Dynamics: Models and Methods, New York: Academic Press, 1984.

Williams, Roberton and Patricia Ruggles. "Determinants of Changes in Income Status and Welfare Program Participation," in Proceedings of the Social and Behavioral Sciences Section, American Statistical Association, December 1987.

Survey of Income and Program Participation

Two Notes on Sampling Variance Estimates from the 1984 SIPP Public-Use Files

8902 79

Barry V. Bye and Salvatore J. Gallicchio
Social Security Administration

April 1989

The views expressed are the authors' and do not necessarily reflect those of the Census Bureau.

Preface

This working paper provides two contributions by Barry Bye and Sal Gallicchio of the Social Security Administration related to the estimation of variances from the SIPP public-use files. The 1984 public-use data files of the Survey of Income and Program Participation provide pseudo stratum and pseudo primary sampling unit codes that permit direct estimates of sampling errors. The first note is a reprint of an October 1988 Social Security Bulletin article describing a methodology for calculating sampling errors directly from the SIPP public-use file. The authors applied this method to the calculation of variances for persons participating in programs administered by the Social Security Administration, and empirically show an apparent sensitivity of generalized variances (as found in the SIPP Users' Guide and Technical Documentation) to curve fitting procedures.

The second note in this working paper reports the results of comparisons of direct variance estimates from the public-use file with variance estimates based on the original sample design (computed by Census Bureau staff). The authors conclude that the variance estimates are very much alike, suggesting some validity for the direct variance estimates using the pseudo design codes.

TABLE OF CONTENTS

Methodology..... 2

Balanced Half-Sample Replication..... 2
Variance Curve..... 3
Generalized Variances for Counts and Proportions..... 4
Variances of Medians..... 4
Statistical Tests for Differences of Medians..... 5

Results..... 5

Participants Aged 18 or Older..... 5
Participants Under Age 18..... 5
Comparison with Census Generalized Variances..... 6
Medians..... 7

Conclusion..... 11

Appendix..... 12

NOTE 2..... 19

Introduction..... 20

Variance Items..... 20

Results..... 22

REFERENCES

Two Notes on Sampling Variance Estimates from the 1984 SIPP Public-Use Files

by Barry V. Bye and Salvatore J. Gallicchio*

The Census Bureau's Survey of Income and Program Participation (SIPP) provides data that can be used to study the characteristics of Old-Age, Survivors, and Disability Insurance (OASDI) and Supplemental Security Income (SSI) program participants. It is important that estimates of sampling errors accompany such studies because the estimates may have large sampling errors due to the small number of sample cases available for specific analyses. The generalized sampling variances provided by the Census Bureau did not identify separately either program's participants and, therefore, do not pertain directly to analyses of these groups. This article describes an approach to the direct computation of sampling variances for OASDI and SSI program participants. The approach uses the pseudo stratum and half-sample codes available in SIPP public use data files. A table of generalized standard errors is constructed for participants of both programs aged 18 or older. Generalized standard errors could not be computed for child beneficiaries under age 18 because of a wide variation of design effects across subpopulation estimates.

The Survey of Income and Program Participation (SIPP) provides data that can be used to study the socioeconomic characteristics of persons participating in programs administered by the Social Security Administration (SSA): Old-Age, Survivors, and Disability Insurance (OASDI) and Supplemental Security Income (SSI).[1] Currently, data from the initial 1984 SIPP panel are available. The 1984 panel consists of approximately 20,000 households comprising about 54,000 individuals. Through a special algorithm developed by SSA, about 8,000 of these individuals have been identified as OASDI and SSI program participants.[2] Included among them are about 4,600 retired-worker beneficiaries, about 600 disabled-worker beneficiaries, and 700 aged, blind, or disabled SSI recipients. The remaining participants are survivor, spouse, or child beneficiaries.

To provide summary SIPP data on SSA program participants to the public, a special set of tables was introduced in the Annual Statistical Supplement to the Social Security Bulletin for 1987.[3] The tables pertain to the civilian noninstitutionalized population receiving OASDI and SSI payments. They focus on three major themes: the composition and level of income of persons receiving different types of OASDI benefits, the general characteristics of persons aged 18-64 receiving OASDI or SSI payments based on disability, and similar information about SSI recipients aged 18 or older. The unit of analysis in these tables is the individual recipient.

Many of the distributions and income levels shown in the Supplement tables are based on a relatively small

*Office of Research and Statistics, Office of Policy, Social Security Administration.

[1] General information on the SIPP can be found in Dawn Nelson, David McMillen, and Daniel Kasprzyk, An Overview of the Survey of Income and Program Participation (SIPP Working Paper Series, No. 8401, update 1), Bureau of the Census, Department of Commerce, 1985.

[2] Denton R. Vaughan, A Survey-Based Type of Benefit Code for the Social Security Program (ORS Working Paper Series), Office of Research and Statistics, Social Security Administration (forthcoming).

[3] Annual Statistical Supplement to the Social Security Bulletin, 1987, Office of Research and Statistics, Social Security Administration, 1987, tables 15-22.

number of sample cases. Summary statistics generated from small numbers of cases can be imprecise due to large sampling errors (variances) and often suggest differences between subpopulations when no real differences exist. It is important, therefore, that estimates of sampling errors be provided along with the estimates of direct interest.

The Bureau of the Census has provided generalized variance curves for a number of quantities from the 1984 SIPP panel.[4] These curves do not identify OASDI or SSI recipients separately; therefore, the curves do not pertain directly to SSA program participants. Fortunately, provisions were made for the direct calculation of sampling variances of SIPP estimates using special codes available in the SIPP public use data files. These codes allocate the SIPP sample cases to a set of pseudo strata and pseudo primary sampling units. The codes permit direct estimates of sampling variances to be obtained by a number of methods.

The results of direct sampling variance computations for SSA program participants are presented in this article. The approach used to estimate the variances was the method of balanced half-sample replication.[5] The appendix at the end of the article includes the detailed specifications for estimating sampling variances from the SIPP using the same techniques that were used for the computations presented in this article. The results of the calculations also are provided in sufficient detail to be used as a benchmark.

Sampling variances were computed for more than 300 population estimates, cross-classifying the recipients by sex, age, marital status, and type of beneficiary. A curve was fit to the estimated variances and was used to produce tables of generalized standard errors. The tables of generalized standard errors can be applied directly to the data presented in the Supplement for program participants aged 18 or older and also can be used with other analyses from wave 1 of the 1984 SIPP panel that pertain to SSA program participation of adults. A separate analysis for child beneficiaries under age 18 showed that estimated standard errors were strongly associated with family size. As a result, tables of generalized standard errors that would be applicable to a variety of estimates for this subpopulation could not be developed.

The generalized variance curve presented in this article yields variance estimates that are markedly different from those generated by curves from the Census Bureau. In part, the difference may be due to the fact that variances of individual items estimated from the pseudo sample design may differ from those estimated directly from the original design. However, a part of the difference appears to be due to differences in the fit of the curves employed by the Census Bureau and by SSA staff, even though the functional form was the same. The SSA results appear to be more appropriate for variance estimates of OASDI and SSI program participants.

Sampling variances were also computed for some of the median income amounts shown in the Supplement. The variances and estimated sampling covariances between the medians were used to test hypotheses about differences in the size of the estimated median income amounts among various subpopulations.

Methodology

Balanced Half-Sample Replication

The method of balanced half-sample replication is an approach to the estimation of sampling variances for complex sample designs that can be implemented easily and has been applied to a wide variety of statistical estimates. This method presupposes that the primary sampling units for the population have been assigned to one of L strata, and two of the units are selected with replacement from each stratum with probability proportionate to size. Half-sample replicates of this design can be formed by selecting one of the two units from each stratum. For a sample design with L strata, there are 2^L such half samples. If an estimate of the statistic of interest is made in each half sample and in the full sample, then the average squared difference between half-sample and full-sample estimates from any subset of half samples provides an estimate of the sampling variance of the statistic. The estimate of the sampling variance is most precise when all 2^L half samples are employed.

When L is large, one would like to use only a part of the 2^L half samples to estimate the sampling variances without loss of precision. It turns out that special sets of half samples, called balanced, orthogonal sets, are particularly good candidates. Estimates of sampling variances from these special sets are algebraically equivalent to those obtained using all half samples. Also, when the full-sample estimate is a linear function of the half-sample estimates, the average estimate over the balanced, orthogonal set will be equal to the full-sample estimate. The minimum number of half samples required for a fully balanced orthogonal set is the smallest multiple of 4 which is greater than the number of strata in the sample design. For designs with many strata, this number will be much smaller than the total number of

[4] Survey of Income and Program Participation, User's Guide, Bureau of the Census, Department of Commerce, July 1987, pages 7-1 through 7-27.

[5] Kirk Wolter, Introduction to Variance Estimation, Springer-Verlag, New York, 1985.

possible half samples. Descriptions of balanced, orthogonal sets for many designs are provided in the literature.[6]

Once a set of half samples has been identified, estimated sampling variances are particularly easy to compute. Let $\theta_\alpha (\alpha = 1, \ldots, K)$ denote the estimator of the population parameter of interest computed from the α th half sample, and let θ be the corresponding estimate from the full sample. An estimator of the sampling variance of θ, $V(\theta)$, based on K half samples is given by

$$V(\theta) = \sum_{\alpha=1}^{K} (\theta_\alpha - \theta)^2/K \qquad (1)$$

When θ is a linear function of the θ_α, so that

$$\theta = \bar{\theta} = \sum_{\alpha=1}^{K} \theta_\alpha/K ,$$

then (1) provides an unbiased estimate of the variance of θ. When θ is not linear in θ_α (for example, θ is a ratio, a median, a correlation coefficient), then $\theta \neq \bar{\theta}$ and the expected value of $V(\theta)$ differs from the variance of θ by an amount often well approximated by $[E(\bar{\theta} - \theta)]^2$. Thus if $\bar{\theta}$ is close to θ, equation (1) will provide a good approximation of the sampling variance when θ is not linear.[7]

Variance Curve

A two-parameter curve was fit to the variance estimates obtained by the replication method. The curve specified the relative variance (Rv), the variance divided by the square of the estimate, as a function of the estimate.

$$Rv(x) = a + b/x \qquad (2)$$

where

a and b are coefficients to be estimated,
x is the estimated population total, and
Rv(x) is the estimated relative variance of x— that is,

$$Rv(x) = V(x)/x^2 .$$

This functional form has provided a fairly good representation of the relationship between Rv(x) and x in other surveys. Its use is motivated by the following considerations.[8]

The design effect (Deff) for a particular estimate, x, from a complex sample design is defined as the ratio of the sampling variance of x under the design to the sampling variance that would have been obtained from a simple random sample of equal size. For a sample of size n from a population of size N, the simple random sampling variance of an estimated total, x is given by

$$var(x) = var(pN) = N^2PQ/n$$

where

P = X/N, is the true population proportion,
X is the population total estimated by x,
Q = 1-P, and
p is the sample estimate of P.

The variance of x from a complex design of the same size can be expressed as

$$var_c(x) = Deff\ (var(x)) = Deff\ (N^2PQ/n).$$

The relative variance of x is given by

$$Rv(x) = var_c(x)/X^2 = Deff\ (Q/Pn)$$

$$= -\ Deff/n + (N/n)Deff/X. \qquad (3)$$

Equation (3) has the same form as equation (2) where a = -Deff/n and b = (N/n)Deff. If it is reasonable to assume that a constant design effect exists for a particular set of estimates, then the estimated relative variances for those items may be accurately represented by a two-term curve of the form in (2) from which generalized variances can be computed.

The method used to estimate the coefficients in (2) was an iterative procedure that minimized the function

$$\sum_{i=1}^{I} \left[\frac{Rv_i - \hat{Rv}_i}{\hat{Rv}_i^*} \right]^2$$

where

Rv_i is the computed relative variance for the ith item;
$\hat{Rv}_i$ is the estimated relative variance for the ith item from the curve;

[6] R. L. Plackett and J. P. Burman, "The Design of Optimum Multifactor Experiments," Biometrika, 33(1946), pages 305 and 325.

[7] Wolter (1985), op. cit., references a number of empirical investigations supporting the use of equation (1).

[8] See, for example, The Current Population Survey: Design and Methodology (Technical Paper 40), Bureau of the Census, Department of Commerce, January 1978.

$\hat{Rv}_i^*$ is a weight for the ith item. It is set equal to the computed relative variance, Rv_i, in the first iteration; for all subsequent iterations it is set equal to the estimated relative variance, $\hat{Rv}_i$, from the previous iteration.

I is the number of items to be fit.

This estimation approach gives greater weight to items with smaller estimated relative variances (and, thus, generally larger estimated totals) and has been found to work well in other surveys.[9]

Generalized Variances for Counts and Proportions

Having estimated values for the coefficients in equation (2), the relative variance for a specific estimated total, x_0, can be obtained by substituting x_0 into that equation. The variance of the estimated total can be obtained by multiplying the relative variance by the square of the estimate.

$$\hat{V}(x_0) = \hat{Rv}(x_0)x_0^2$$

$$= ax_0^2 + bx_0 \qquad (4)$$

Equation (4) can also be used to produce generalized estimates of variances of proportions. A proportion is the ratio of two estimated totals, $p = x/y$, where the cases counted in the numerator are a subset of the cases counted in the denominator. In large samples, the relative variance of this type of ratio can be approximated by the following formula:

$$Rv(p) = Rv(x/y) = Rv(x) - Rv(y)$$

or

$$V(p) = V(x/y) = (x/y)^2 [Rv(x) - Rv(y)] \qquad (5)$$

Substitution of estimates from (2) into (5) provides generalized variance estimates for proportions.

$$\hat{V}(p) = p^2[b(1/x - 1/y)] = (b/y)(p)(1 - p) . \qquad (6)$$

Tables of generalized standard errors for estimated totals are often produced from equation (4) by computing and displaying the square root of the estimated variances for a set of predetermined values of x. Similarly, a table of standard errors for estimated proportions can be computed from (6). This table will be two dimensional with the size of the base of the percent on one dimension and the estimated proportion on the other.

Variances of Medians

The balanced half-sample replication approach was used to estimate standard errors for the estimated medians in table 17 of the 1987 Supplement. That table presents median OASDI income, median total income, and the median of the ratio of OASDI income to total income for several beneficiary groups, cross-classified by a number of factors.

In this article, the medians were estimated from distributions of the variables of interest using the following formula:[10]

$$M = L_j + \left[\frac{S_{50} - S_j}{N_j}\right] W_j$$

where

j indexes the interval containing the 50th percentile;

L_j is the lower limit of the jth interval;

S_{50} is the estimated population at the 50th percentile;

S_j is the estimated population with values below the jth interval;

N_j is the estimated population in the jth interval; and

W_j is the width of the jth interval.

An interval width of $25 was used for the OASDI income distribution. Intervals of $50 or $100 were employed for the total income distribution, the latter used to capture the larger monthly benefit amounts. An interval of .05 was used for the income ratio.

The sampling variance of M was obtained by estimating M in each half sample and then applying

[9] There is no specific justification for this weighted least squares approach other than the usefulness of its results. Ordinary least squares estimates, minimizing

$$\sum_{i=1}^{I} (Rv_i - \hat{Rv}_i)^2 ,$$

have been found to give too much weight to small estimates, x, with characteristically large estimated relative variances. Nonlinear least squares estimates, minimizing

$$\sum_{i=1}^{N} \left[\frac{Rv_i - \hat{Rv}_i}{\hat{Rv}_i}\right]^2 ,$$

appear to give too much weight to observations with large estimated totals.

[10] The estimated medians shown in the Supplement were computed by the TPL tabulation program on an IBM system. The medians reported here were computed by the PASS tabulation program on a UNIVAC system and they sometimes differ from the Supplement estimates by small amounts.

equation (1). This approach was repeated for each of the three median amounts and for each subpopulation.

Statistical Tests for Differences of Medians

Statistical tests were made on the variation in medians across the categories of a particular variable (sex, age, and size of family, for example) within a particular beneficiary group. The test approach follows that developed by Grizzle, Starmer, and Koch.[11] Let $M_1, M_2, \ldots, M_k$ be a set of estimated medians for k categories of the variable. Then a χ^2— type test statistic for the hypothesis $H_0 : M_1 = M_2 = \ldots = M_k$ can be constructed under the assumptions that the M have, jointly, a multivariate normal distribution and that a consistent estimate of the sampling covariance matrix is available.[12]

The sampling covariance matrix is obtained through the balanced half-sample method by a computation similar to that of equation (1). The (i,j)th element of the matrix is given by

$$\sum_{\alpha=1}^{K} [M_\alpha^{(i)} - M^{(i)}]\,[M_\alpha^{(j)} - M^{(j)}]/K \,.$$

where

$M^{(r)}$ is the estimate of the median for the rth category from the entire population,

$M_\alpha^{(r)}$ is the estimate of the median for the rth category from the α th half sample, and

K is the number of half samples.

Among retired-worker beneficiaries, in two cases, the set of categories consists of a cross-classification of two factors: sex by age and sex by marital status. In these cases, a sex effect, an age (or marital status) effect and a combined effect were tested. For disabled-worker beneficiaries, the type-of-family categories refer to both marital status and presence of minor children. In this case, the medians for married versus not married and the medians for married with minor children versus married with no minor children were tested.

[11] J. R. Grizzle, C. F. Starmer, and G. C. Koch, "Analysis of Categorical Data by Linear Models," Biometrics, September 1969, pages 489-504.

[12] The asymptotic normality of the estimated medians follows from the asymptotic normality of the estimated ratios $(S_{50}/N_j, S_i/N_j)$ of which the median is a linear function. The covariance matrix computed by half-sample replication on the pseudo design is not a consistent estimate. Still, it is believed that the GSK test statistics provide useful information about the real spread in the medians, even if the true significance levels are not known.

Results

Participants Aged 18 or Older

Appendix table I presents the population estimates, standard errors, and relative variances for each of the items described above. There were 326 subpopulation estimates based on more than 1 sample case. The estimates ranged from a low of about 7,000 based on 2 sample cases to a high of 38 million based on 7,943 sample cases that represent the entire OASDI and SSI recipient population.[13] The variance curve that was dervied from the items has coefficients[14]

$$a = .0007$$
$$b = 5217.$$

Tables of generalized standard errors based on this curve follow.[15] For the estimated totals of a specific size, table 1 gives one standard error of the estimate. Table 2 gives one standard error for estimated proportions with bases of various sizes.

Participants Under Age 18

When constructing estimates of family characteristics for children, one would expect large design effects in the estimated sampling errors. All children will tend to report (or have coded for them) the same family data, thus reducing the effective number of independent observations by the average number of children per family. Because OASDI benefits awarded to minor children tend to be divided among all the children in a beneficiary family, the strong clustering effects that one finds for child-related estimates are expected to appear for beneficiary children as well.

To investigate the sampling variances for children, a set of estimates was constructed by cross-classifying

[13] A sampling variance cannot be estimated for totals based on 1 sample case. Algebraically, the balanced half-sample estimator yields a perfect 1.0 for the estimated relative variance. Thirty-nine of these cells are shown in appendix table I.

[14] The estimated constant, a, is positive. Although the rationale presented suggests that a should be negative, the algorithm used to estimate the parameters does not impose this constraint. The estimated design effect from the b coefficient is

$$\text{Deff} = b(n/N) = (5217)\,(7943/34160810) = 1.2.$$

Values for n and N are obtained from the first item in the variance table in the appendix.

[15] Variance curves were also estimated for sets of items for several subpopulations of the total beneficiary population: disabled workers, persons aged 65 or older, and persons receiving SSI payments. Generally, the sizes of standard errors for similar size cells across these groups did not differ. A curve was also estimated for the group aged 18 or older, using items derived from cross-classifying age, family size, and family income. Again, no substantial differences were seen in estimated a and b parameters.

family size, family income, sex, and race. As expected, a variance curve fit to all of the items exhibited a systematic lack of fit, overestimating the computed variances for smaller families and underestimating the variances for larger families. Fitting separate curves by family size resulted in the following set of a and b parameters:

Family size	Parameter a	Parameter b
1-3	.0034	4922.
4	.0127	5849.
5 or more	.0199	8733.

The increasing values of both the a and b parameters indicate that substantial increases in sampling variances are to be expected, for an estimate of fixed size, as family size increases.

Table 1.—Standard errors for estimated population totals

Estimate	Standard error
25,000	11,436
50,000	16,202
75,000	19,878
100,000	22,994
250,000	36,738
500,000	52,842
750,000	65,786
1,000,000	77,176
2,500,000	132,954
5,000,000	211,284
7,500,000	284,417
10,000,000	355,574
25,000,000	771,177
50,000,000	1,455,403

These results imply that the sampling variance for an estimated subpopulation of child beneficiaries under age 18 will depend largely on the family size composition of the subpopulation. A set of child-beneficiary estimates would not be likely to exhibit a constant design effect; and therefore, it is unlikely that a two-term curve of the kind described above would provide a good approximation to the estimated sampling variances for the set. Accordingly, no generalized variances for child beneficiaries are presented. There appears to be no substitute for direct variance calculations in this case.

Comparison with Census Generalized Variances

The SIPP User's Guide presents parameters for a number of generalized variance curves[16] From the descriptions associated with the various Bureau of the Census curves, one might suppose that curve 1, "program participation and benefits, poverty," would be the appropriate curve for OASDI and SSI program participation estimates. Because the generalized variances computed from the pseudo design differ so greatly from those obtained from Census curve 1, some discussion is needed.

Table 3 shows estimated standard errors from the SSA curve and Census curve 1 for a range of estimates.[17] For estimates less than 10 million, the Census estimates are 1.20 to 1.75 times larger than those from the SSA curve. Some of this difference could be due to differences in computational schemes for the direct

[16] SIPP User's Guide, op. cit., page 7-5.

[17] The parameters from Census curve 1 are:

a = -.0000942, and b = 16059.

Table 2.—Standard errors for estimated percents

Base of percent	Percent 1 or 99	2 or 98	5 or 95	8 or 92	10 or 90	15 or 85	20 or 80	25 or 75	30 or 70	35 or 65	40 or 60	50
25,000	4.54	6.39	9.95	12.39	13.70	16.31	18.27	19.77	20.93	21.78	22.37	22.83
50,000	3.21	4.52	7.04	8.76	9.69	11.53	12.92	13.98	14.80	15.40	15.02	16.14
75,000	2.62	3.69	5.75	7.15	7.91	9.41	10.55	11.42	12.08	12.58	12.92	13.18
100,000	2.27	3.20	4.98	6.19	6.85	8.15	9.13	9.89	10.46	10.89	11.19	11.42
250,000	1.44	2.02	3.15	3.92	4.33	5.16	5.78	6.25	6.62	6.89	7.07	7.22
500,000	1.02	1.43	2.23	2.77	3.06	3.65	4.00	4.42	4.68	4.87	5.00	5.11
750,000	.83	1.17	1.82	2.26	2.50	2.98	3.33	3.61	3.82	3.98	4.08	4.17
1,000,000	.72	1.01	1.57	1.96	2.17	2.58	2.89	3.13	3.31	3.44	3.54	3.61
2,500,000	.45	.64	1.00	1.24	1.37	1.63	1.83	1.98	2.09	2.18	2.24	2.20
5,000,000	.32	.45	.70	.88	.97	1.15	1.29	1.40	1.48	1.54	1.58	1.61
7,500,000	.26	.37	.57	.72	.79	.94	1.05	1.14	1.21	1.26	1.29	1.32
10,000,000	.23	.32	.50	.62	.68	.82	.91	.99	1.05	1.09	1.12	1.14
25,000,000	.14	.20	.31	.39	.43	.52	.58	.63	.66	.69	.71	.72
50,000,000	.10	.14	.22	.28	.31	.36	.41	.44	.47	.49	.50	.51

Table 3.—Comparison of generalized standard errors for estimated totals

Estimate	SSA	Census curve 1	Percent
25,000	11440	20035	175.1
50,000	16206	28332	174.8
75,000	19882	34697	174.5
100,000	22997	40062	174.2
250,000	36731	63316	172.4
50,0000	52805	89476	169.4
750,000	65708	109505	166.7
100,000	77051	126352	164.0
250,000	132446	198894	150.2
500,000	209962	279177	133.0
750,000	282181	339328	120.3
10,000,000	352375	388806	110.3
25,000,000	761853	585320	76.8

variance estimates on which the curves are based. Both the variance estimators and the assumed sample design are different.[18]

Much of the difference in the curves, however, appears to be attributable to differences in curve-fitting strategies. The Census curve is based on 36 estimated totals for persons aged 16 or older involving receipt of cash and noncash benefits and labor-force activity. Thirteen of the 36 items are estimates of the Hispanic population with selected characteristics. Unpublished Census Bureau data suggest that variances from curve 1 for population totals of less than 500,000 are substantially overestimated.[19] This is not surprising because only several observations are in this range among the 36 items and they are given little weight by the kind of curve-fitting algorithm described above.[20] As indicated in the appendix, the set of items from which the SSA curve was derived contains a large number of small estimates. The SSA curve appears to fit the observations well for small estimated totals.

The reasons for differences between Census Bureau and SSA curves for larger estimates are more difficult to discern. There is some indication that the design effects for the Hispanic population estimates are larger than those for the corresponding estimates for all races combined, raising the overall level of the Census curve. It is also possible that the design effects for adult OASDI and SSI program participants are generally smaller than the effects for the Census items. Less clustering may occur among OASDI and SSI adult recipients in families and households, compared with recipients in other transfer programs. The small number of items on which the Census curve is based makes a more detailed analysis difficult. At this point, the SSA curve appears to be much preferred for OASDI and SSI program participation estimates.

Medians

The standard errors for the medians in table 17 of the Annual Statistical Supplement are shown in table 4. With the exception of child beneficiaries, the variances of the estimated medians appear to be quite small. The sizes of the estimated standard errors rarely exceed 10 percent of the corresponding medians and are often well under 5 percent. The median income amounts for families of child beneficiaries show larger standard errors than, for example, similar estimates for families of disabled-worker beneficiaries even when the unweighted case counts are about the same. The larger estimated standard errors are probably the result of the clustering effects for child beneficiaries discussed above.

The generally small standard errors are also reflected in the test statistics for the hypotheses concerning differences of medians. For each set of categories and each type of median, the differences between medians across categories were statistically significant at the .05 level in most cases. When contrasts were significant, the significance levels tended to be much smaller than .05, usually less than .0001.

The contrasts that were not significant at the .05 level are described at the end of table 4. The table identifies the specific comparisons and provides the value of the test statistic, the degrees of freedom, and the p-value. The following examples demonstrate how the test results can be interpreted.

The statistical tests indicated no two-way interaction existed between sex and age regarding the ratio of OASDI benefits to total income for retired-worker beneficiaries. Differences in median ratios between age groups tended to be about the same for both men and women. The differences between median ratios for men by age group are 13, 9, and 0. The corresponding differences for women are very similar (12, 7, and 2). The statistical tests did show significant sex differences and significant age differences. The pattern of median ratios, therefore, can be described by adding sex and age effects without the need to adjust for particular sex-age combinations.

[18] Census estimates were computed by the half-sample replication method using a set of 50 half samples that was not fully balanced. The appendix provides a brief description of the procedures used to create the pseudo design codes.

[19] For a description of the items, see "Memorandum for Documentation from Karen E. King, Subject: SIPP Variances: Items by Generalized Variance Parameter," Bureau of the Census, Department of Commerce, June 19, 1985. The Census direct variance estimates are unpublished and were made available by the Statistical Methods Division, Bureau of the Census.

[20] The Census Bureau curve-fitting algorithm differed from that described above in that the relative variance for the overall population total, T, was constrained to be zero. Thus, $a + b/T = 0$ or $a = -b/T$, and b is estimated from a one parameter model $V(x) = b(1/x - 1/T)$. This approach is reasonable because the case weights are adjusted to achieve certain population totals. However, imposing this constraint may also contribute to the overestimate of the variance for small population estimates.

Table 4.—Standard errors for table 17, Annual Statistical Supplement to the Social Security Bulletin, 1987

Characteristic	OASDI benefit		Total income		Ratio*	
	Median	Standard error	Median	Standard error	Median	Standard error
	Retired workers					
Total	577	10	1210	23	53	1
Men	633	10	1300	30	51	1
Women	515	7	1096	29	57	1
Sex and age of beneficiary:[1]						
Men—						
62-64	502	11	1442	54	34	2
65-69	672	18	1444	51	47	2
70-74	682	13	1282	40	56	2
75 or older	611	16	1137	35	56	1
Women—						
62-64	582	39	1481	76	41	2
65-69	569	19	1216	28	53	2
70-74	531	12	1072	42	60	2
75 or older	469	9	847	45	62	2
Sex and marital status:						
Men—						
Married	697	9	1417	26	50	1
Widowed	456	13	946	64	49	2
Divorced	451	33	759	93	64	4
Never married	476	34	893	79	56	3
Women—						
Married	763	8	1487	38	52	2
Widowed	437	6	760	28	61	2
Divorced	411	13	778	57	58	4
Never married	452	20	935	115	58	3
Size of family:						
1 person	419	6	629	19	65	1
2 persons	713	9	1351	28	54	1
3 persons or more	669	29	2261	74	30	1
Monthly family income:						
Less than $500	326	7	396	6	91	1
$500-$999	520	5	743	7	74	1
$1,000-$1,499	713	15	1225	7	57	1
$1,500-$1,999	718	15	1722	14	41	1
$2,000-$2,499	793	13	2203	13	35	1
$2,500-$2,999	710	41	2776	20	25	1
$3,000 or more	764	29	3891	83	17	1
Family source of income:						
Earnings[2]—						
Yes	572	15	1946	36	31	1
No	580	13	1015	29	63	1
Assets—						
Yes	622	9	1337	26	50	1
No	428	11	604	24	75	2
Means-tested cash benefits—						
Yes	335	16	594	56	58	1
No	600	9	1247	20	53	1
Other cash income—						
Yes	651	11	1461	23	46	
No	497	7	795	24	71	2
	Disabled workers					
Total[3]	522	14	1162	47	49	2
Men[3]	566	12	1175	57	50	3
Women	419	26	1137	59	46	4
Age of beneficiary:[4]						
18-54	544	16	1240	83	45	4
55-64	501	18	1127	53	50	3

See footnotes at end of table.

Table 4.—Standard errors for table 17, Annual Statistical Supplement to the Social Security Bulletin, 1987—Continued

Characteristic	OASDI benefit		Total income		Ratio*	
	Median	Standard error	Median	Standard error	Median	Standard error
	Disabled workers—cont.					
Size of family:						
1 person	392	26	490	39	79	5
2 persons	547	21	1202	51	44	3
3 persons or more	597	25	1625	162	39	3
Type of family:[5]						
Married	578	15	1367	97	44	2
With minor children	713	48	1284	125	54	6
No minor children	547	17	1427	115	41	3
Unmarried	434	21	833	50	55	5
Monthly family income:						
Less than $1,000	437	19	620	42	80	3
$1,000-$1,999	616	20	1369	49	44	2
$2,000 or more	563	43	2664	113	18	1
Family source of income:						
Earnings[6]—						
Yes	516	17	1831	69	31	2
No	528	20	803	50	70	3
Assets—						
Yes	566	23	1512	90	41	2
No	483	16	822	53	63	4
Means-tested cash benefits—						
Yes	407	34	858	67	52	4
No	553	16	1266	65	47	3
Other cash income—						
Yes	594	20	1574	75	41	2
No	477	14	884	48	62	5
	Nondisabled widows					
Total	379	8	657	33	59	2
Age of beneficiary:[7]						
60-69	363	12	834	43	47	3
70 or older	386	9	579	25	68	3
Size of family:						
1 person	363	10	471	18	72	2
2 persons	458	19	1227	82	41	5
3 persons or more	373	15	2104	210	17	2
Monthly family income:						
Less than $1,000	361	9	478	10	79	2
$1,000-$1,999	443	21	1304	36	32	2
$2,000 or more	401	16	2939	84	13	1
Family source of income:						
Earnings[8]—						
Yes	368	10	1759	184	19	2
No	385	10	496	20	75	2
Assets—						
Yes	403	7	825	38	51	2
No	316	11	405	15	81	3
Means-tested cash benefits—						
Yes[9]	258	12	454	32	59	3
No	396	7	706	34	59	3
Other cash income—						
Yes	406	16	1033	69	39	2
No	369	8	525	21	72	2

See footnotes at end of table.

Table 4.—Standard errors for table 17, Annual Statistical Supplement to the Social Security Bulletin, 1987—Continued

Characteristic	OASDI benefit		Total income		Ratio*	
	Median	Standard error	Median	Standard error	Median	Standard error
	Minor children					
Total	604	41	1463	114	43	3
Size of family:[10]						
1 or 2 persons	392	61	981	132	43	11
3 persons	622	77	1437	155	50	7
4 persons	674	69	1578	252	46	10
5 persons	543	101	1800	198	30	5
6 persons or more	539	90	1345	213	45	5
Type of family:[11]						
With husband/wife head	601	42	1828	112	32	3
With single head	615	75	1181	70	49	5
Monthly family income:						
Less than $1,000	464	33	674	57	81	5
$1,000-$1,999	700	48	1449	79	46	3
$2,000 or more	675	89	2928	189	20	3
Family source of income:						
Earnings—						
Yes	519	34	1829	78	31	2
No[12]	728	61	958	48	86	6
Assets[13]—						
Yes	655	53	1999	99	30	3
No	525	43	973	66	70	5
Means-tested cash benefits[15]—						
Yes	454	42	966	150	56	8
No	657	35	1713	133	39	3
Other cash income[14]—						
Yes	645	56	1911	66	34	3
No	541	50	1251	86	49	7

* OASDI divided by total; two decimals implied.

Finding	Chi**2	d.f.	p-value
[1]No two-way interaction in ratio	1.25	3	.74
[2]No difference in OASDI benefit level	.50	1	.70
[3]No difference in total income	.27	1	.60
No difference in ratio	.90	1	.34
[4]No difference in total income	1.56	1	.21
[5]No difference in total income for married with minor/with no minor	.69	1	.41
No difference in ratio for married with minor/with no minor	3.58	1	.06
[6]No difference in OASDI benefit level	.22	1	.64
[7]No difference in OASDI benefit level	2.60	1	.11
[8]No difference in OASDI benefit level	1.54	1	.22
[9]No difference in ratio	.02	1	.89
[10]No difference in ratio	7.26	4	.12
[11]No difference in OASDI benefit level	.02	1	.88
[12]No difference in OASDI benefit level	3.02	1	.08
[13]No difference in ratio	3.73	1	.05
[14]No difference in OASDI benefit level	1.56	1	.21

In contrast to the sex-age findings for retired workers, the sex by marital status tests showed that a two-way interaction was required to describe the patterns of median ratios. Again, differences were seen among the medians for each factor separately, but the pattern of marital status differences was not the same for men and women. Note, for example, that the difference in median ratios for married men and widowed men, -1, appears to be quite different from the difference between the medians of married and widowed women, +9. Among the other sequential contrasts differences were also evident. This pattern of values can not be explained by additive effects alone.

Sex and age contrasts for disabled-worker beneficiaries present situations in which a significant difference existed among median OASDI benefits but not among total incomes or ratios. This apparent inconsistency could be due to chance alone. However, there could be another explanation. The median ratio is not, algebraically, the same quantity as the ratio of the medians. It is possible that the ratios of the medians in the population are different, as suggested by the data presented here, but that the median ratios in the population are the same.

The remaining findings of differences in medians generally indicate that a contrast between one pair of medians was not significant. The one exception is the contrast of family size ratios for families with minor children. Because there were five family size categories, four contrasts were involved in the comparison.

Conclusion

This article described a methodology for calculating sampling errors directly from the SIPP public use file and applied this method to the calculation of variances for persons participating in SSA-administered programs. The methodology is presented in sufficient detail so that researchers can apply the same methods to their specific analyses. Since the replication variance estimation approach is not difficult to implement and facilitates a wide range of hypothesis testing techniques, it is recommended that direct variance calculations be used. This position is further supported by the apparent sensitivity of generalized variances to curve-fitting procedures. Estimating variances directly will also permit variances to be obtained from subsequent waves of the 1984 SIPP panel. Presumably, estimated standard errors will be higher for later waves of the panel due to the accumulated sample attrition at each wave.

For those who cannot compute variances directly, standard error tables have been provided for OASDI and SSI program participants aged 18 or older from wave 1 of the 1984 panel. The standard errors pertain directly to the SIPP tables in the Annual Statistical Supplement to the Social Security Bulletin for 1987. The standard error tables can also be used for other analyses of program participants from wave 1. Generalized standard errors for participants under age 18 could not be developed.

Several matters need further investigation to raise confidence in direct sampling error estimates from the public use files. A comparison of variance estimates from the pseudo design and from the actual sample design will show whether the pseudo design yields estimates that are, on average, smaller than those obtained when the original design is used. A comparison of the size of test statistics of the type that are used in this article also would be useful. These statistics require estimates of sampling variances and covariances, and it would be helpful to know if the pseudo design yields reasonable estimates of covariance as well as variance. Finally, little is known about the raw sample sizes required before normality is achieved in the sampling distribution of the various statistics presented. If for small samples the sampling distribution of counts, proportions, or medians is markedly different from normal, it might be misleading to form confidence intervals or to perform statistical tests assuming a normal distribution (that is, assuming symmetric intervals of 1 standard error about the estimate yields a 68-percent confidence interval, 2 standard errors provides a 95-percent confidence interval). The true confidence intervals may be larger or smaller than those of a normal distribution and may not be symmetric about the estimate. All of these matters are important if the Survey of Income and Program Participation is to be used for making inferences about the population under SSA-administered programs and not just for descriptive reporting.

Appendix: Detailed Sampling Variance Specifications

Assignment of Half-Sample Codes

Each person in the sample in the 1984 SIPP public use file had been assigned a pseudo-stratum code and a pseudo primary sampling unit (PSU) code within each pseudo stratum.[1] Generally, a self-representing (SR) PSU from the original design was associated with two non-self-representing (NSR) PSUs to form a pseudo stratum. Segments of the SR PSU were assigned to one of the two pseudo PSUs at random; each of the NSR PSUs was assigned, in its entirety, to one or the other of the pseudo units. In some cases, two SR PSUs or four NSR PSUs were grouped to form a pseudo stratum. The assignment resulted in the formation of 71 pseudo strata with 2 pseudo units in each stratum. The original PSU codes were withheld from the public use file to prevent access to small geographic areas where a risk of disclosure of individual identities might be possible.

For a design with 71 strata with two units each, the smallest number of half samples that can achieve full orthogonal balance is 72. The set of balanced half samples used in the variance computations is shown in chart 1.[2] The array represents a string of 72 1s and 0s for each of the 71 pseudo strata. For a SIPP sample case in pseudo-stratum δ and pseudo-unit 1, the string in the δth row of the array was attached to the record. For a SIPP sample case in pseudo-stratum δ and pseudo-unit 2, the complement (that is, 1s replaced by 0s, and vice versa) of the string in the δth row of the array was attached. These strings effectively assign each SIPP case to 36 of the 72 half samples. A "1" in the αth position in the string indicates that the case is to be included in the αth half sample; a "0" means that the case is not to be included.

Item Specification for Generalized Variances

Replication variances were obtained for estimated population totals of OASDI and SSI recipients. Recipiency status was determined by the responses for September 1983. Estimated population totals were obtained in each half sample by multiplying the sum of the weights by 2. [3] The recipients were cross-classified by age, sex, marital status, and type of recipient (OASDI only, SSI only, and concurrent OASDI and SSI). This cross-classification yielded 326 distinct detailed and subtotal cells with more than one case.

The September 1983, OASDI and SSI recipient universe consists of those persons in the sample who meet the following test:[4]

$$[(\text{IO1AMT-*} > 0 \text{ or } \text{IO3AMT-*} > 0) \text{ or } (\text{SOCSEC-*} = 1 \text{ and } \text{AGE-*} < 18)] \text{ and } [\text{FNLWGT-*} > 0]$$

where

IO1AMT-*	refers to the OASDI benefit amount;
IO3AMT-*	refers to the SSI amount;
SOCSEC-*	is the OASDI indicator;
AGE-*	is age in September 1983; and
FNLWGT-*	is the case weight.

Each variable is selected for September based on the rotation group of the sample case shown below:

Rotation group	Month
1	4
2	3
3	2
4	1

The cross-classifying variables (type of benefit, age, sex, and marital status) were constructed as follows:

Age (AGE-*):

Under 18	65-69
18-24	70-74
25-34	75-84
35-44	85 or older
55-64	

Sex:

Male, Female

Type of benefit:

OASDI only	(IO1AMT-* > 0 and IO3AMT-* = 0) or (SOCSEC-* = 1 and AGE-* < 18)
SSI only	(IO1AMT-* = 0 and IO3AMT-* > 0)
OASDI and SSI	(IO1AMT-* > 0 and IO3AMT-* > 0)

Marital status (MS-*):

	Code
Married	Under 2
Widowed	3
Separated	4, 5
Never married	6 or over

Table I presents the estimated sampling variances for the 326 items described above.

[1] These fields are identified as H*-STRAT and H*-HSC in the public use file data dictionary. The codes for month 1 were used. The codes do not vary by month.

[2] The 72 order design in Plackett and Burman (1946), op.cit., was used. The array can be generated by shifting the first row one digit to the left for each subsequent row.

[3] This estimator does not fully replicate the original SIPP estimator in each half sample. The original SIPP estimator consisted of a number of multiplicative adjustments to the raw case weights. Similar adjustments should have been applied separately in each half sample to properly replicate the full sample estimator. The overall effect on the estimated variance of not having done this is unknown.

[4] All variables are referred to by their public use file variable names.

Table I.—Variance estimates for OASDI and SSI participants under SSA-administered programs

Age	Sex	Marital status[1]	Unweighted count	Estimate	Standard error	Relative variance
All program participants						
Total	Total	Total	7943	34160810.	883445.	.0006688
Total	Total	NM	1147	4938770.	207858.	.0017713
Total	Total	S	497	2291038.	99936.	.0019027
Total	Total	W	2307	9917379.	305171.	.0009469
Total	Total	M	3992	17013620.	568181.	.0011153
OASDI only						
Total	Total	Total	7242	31012390.	814853.	.0006904
Total	Total	NM	973	4148071.	191974.	.0021419
Total	Total	S	358	1634194.	91508.	.0031356
Total	Total	W	2078	8966302.	277238.	.0009560
Total	Total	M	3833	16263820.	556481.	.0011707
<18	Male	NM	252	1051521.	89736.	.0072828
<18	Female	NM	256	1064085.	87690.	.0067913
<18	Total	NM	508	2115606.	146801.	.0048149
18-24	Male	S	1	46.	5646.	1.0000000
18-24	Male	NM	30	139714.	27131.	.0377100
18-24	Male	Total	31	145360.	28694.	.0389663
18-24	Female	W	3	10502.	6079.	.3350419
18-24	Female	NM	26	112174.	19133.	.0290918
18-24	Female	Total	29	122676.	20793.	.0287286
18-24	Total	NM	56	251888.	34246.	.0184839
18-24	Total	Total	60	268036.	36677.	.0187243
25-34	Male	M	6	29086.	12232.	.1768577
25-34	Male	W	1	4053.	4053.	1.0000000
25-34	Male	S	3	31835.	24101.	.5731619
25-34	Male	NM	16	89563.	23121.	.0666412
25-34	Male	Total	26	154536.	33560.	.0471601
25-34	Female	M	10	47962.	16933.	.1246478
25-34	Female	W	16	71050.	16858.	.0562995
25-34	Female	S	1	4030.	4030.	1.0000000
25-34	Female	NM	12	54016.	19449.	.1296431
25-34	Female	Total	39	177057.	31562.	.0317771
25-34	Total	M	16	77048.	21730.	.0795461
25-34	Total	W	17	75103.	17339.	.0532992
25-34	Total	S	4	35865.	24436.	.4642159
25-34	Total	NM	28	143579.	32466.	.0511296
25-34	Total	Total	65	331593.	42328.	.0162944
35-44	Male	M	14	61855.	15321.	.0613515
35-44	Male	W	1	4392.	4392.	1.0000000
35-44	Male	S	2	8136.	8136.	1.0000000
35-44	Male	NM	9	47179.	16125.	.1168245
35-44	Male	Total	26	121560.	21518.	.0313335
35-44	Female	M	31	136991.	26813.	.0383101
35-44	Female	W	25	105580.	19971.	.0357782
35-44	Female	S	11	49041.	15943.	.1056871
35-44	Female	NM	7	33957.	12997.	.1464932
35-44	Female	Total	74	325569.	43557.	.0178995
35-44	Total	M	45	198846.	30938.	.0242071
35-44	Total	W	26	109972.	20448.	.0345724
35-44	Total	S	13	57176.	17899.	.0979968
35-44	Total	NM	16	81136.	20711.	.0651601
35-44	Total	Total	100	447129.	49484.	.0122478
45-54	Male	M	52	220557.	28133.	.0162699
45-54	Male	W	2	7013.	4964.	.5011174
45-54	Male	S	17	75694.	18987.	.0629197
45-54	Male	NM	12	58138.	17104.	.0865495
45-54	Male	Total	83	361401.	34312.	.0090141
45-54	Female	M	50	210502.	31456.	.0223298
45-54	Female	W	24	102704.	25139.	.0599145
45-54	Female	S	11	46439.	14031.	.0912957
45-54	Female	NM	6	26079.	10685.	.1678766
45-54	Female	Total	91	385723.	37089.	.0092456
45-54	Total	M	102	431059.	48038.	.0124192
45-54	Total	W	26	109717.	26180.	.0569375
45-54	Total	S	13	122132.	23911.	.0383306
45-54	Total	NM	23	84217.	20167.	.0573444

Table I.—Variance estimates for OASDI and SSI participants under SSA-administered programs—Continued

Age	Sex	Marital status[1]	Unweighted count	Estimate	Standard error	Relative variance
			OASDI only—cont.			
45-54	Total	Total	174	747124.	54047.	.0052331
55-64	Male	M	342	1488914.	99257.	.0044441
55-64	Male	W	26	128374.	24778.	.0372551
55-64	Male	S	36	165105.	29969.	.0329479
55-64	Male	NM	17	82124.	21419.	.0680217
55-64	Male	Total	421	1864517.	113389.	.0036984
55-64	Female	M	351	1478573.	93865.	.0040301
55-64	Female	W	202	856463.	63475.	.0054927
55-64	Female	S	41	174779.	28070.	.0257925
55-64	Female	NM	24	103215.	22004.	.0454497
55-64	Female	Total	618	2613029.	120423.	.0021239
55-64	Total	M	693	2967487.	165997.	.0031291
55-64	Total	W	228	984837.	68234.	.0048003
55-64	Total	S	77	339884.	46806.	.0189647
55-64	Total	NM	60	185339.	32915.	.0315395
55-64	Total	Total	1039	4477546.	197917.	.0019538
65-69	Male	M	652	2778693.	145189.	.0027301
65-69	Male	W	38	173900.	31586.	.0329904
65-69	Male	S	42	197829.	30920.	.0244292
65-69	Male	NM	39	178509.	28946.	.0262943
65-69	Male	Total	771	3328931.	158555.	.0022686
65-69	Female	M	603	2445450.	124833.	.0026058
65-69	Female	W	328	1301091.	63726.	.0023989
65-69	Female	S	68	269385.	34190.	.0161081
65-69	Female	NM	53	210263.	35869.	.0291007
65-69	Female	Total	1052	4226188.	146084.	.0011948
65-69	Total	M	1255	5224143.	228339.	.0019104
65-69	Total	W	366	1474991.	73343.	.0024725
65-69	Total	S	110	467214.	48524.	.0107864
65-69	Total	NM	92	388772.	41663.	.0114844
65-69	Total	Total	1823	7555119.	246535.	.0010648
70-74	Male	M	526	2211887.	125904.	.0032400
70-74	Male	W	69	308203.	45817.	.0220994
70-74	Male	S	28	121108.	23433.	.0374377
70-74	Male	NM	27	125257.	24585.	.0385257
70-74	Male	Total	650	2766455.	139422.	.0025399
70-74	Female	M	377	1634980.	104934.	.0041192
70-74	Female	W	379	1626694.	88937.	.0029892
70-74	Female	S	37	162834.	31180.	.0366651
70-74	Female	NM	46	209242.	34337.	.0269301
70-74	Female	Total	839	3633749.	178731.	.0024193
70-74	Total	M	903	3846867.	199390.	.0026865
70-74	Total	W	448	1934897.	107103.	.0030640
70-74	Total	S	65	283942.	37106.	.0170774
70-74	Total	NM	73	334499.	47244.	.0199480
70-74	Total	Total	1489	6400204.	267776.	.0017505
75-84	Male	M	468	1988365.	125679.	.0039952
75-84	Male	W	116	510172.	61289.	.0144324
75-84	Male	S	28	116411.	24034.	.0426257
75-84	Male	NM	22	95184.	15865.	.0277809
75-84	Male	Total	634	2710130.	150989.	.0031039
75-84	Female	M	269	1191177.	84073.	.0049815
75-84	Female	W	585	2679240.	132442.	.0024436
75-84	Female	S	36	160437.	28486.	.0315242
75-84	Female	NM	88	397776.	47085.	.0140117
75-84	Female	Total	978	4428629.	174050.	.0015446
75-84	Total	M	737	3179542.	190234.	.0035797
75-84	Total	W	701	3189411.	153949.	.0023299
75-84	Total	S	64	276848.	36552.	.0174319
75-84	Total	NM	110	492959.	50716.	.0105844
75-84	Total	Total	1612	7138760.	283838.	.0015809
85+	Male	M	57	246861.	32533.	.0173675
85+	Male	W	44	242744.	42750.	.0310149
85+	Male	S	4	18399.	9514.	.2673954
85+	Male	NM	6	35978.	15424.	.1838019
85+	Male	Total	111	543980.	58333.	.0114989
85+	Female	M	25	91970.	17962.	.0381441
85+	Female	W	219	834132.	63100.	.0057225

Table I.—Variance estimates for OASDI and SSI participants under SSA-administered programs—Continued

Age	Sex	Marital status[1]	Unweighted count	Estimate	Standard error	Relative variance
			OASDI only—cont.			
85+	Female	S	7	27090.	10365.	.1463952
85+	Female	NM	10	34102.	10917.	.1024837
85+	Female	Total	261	987293.	71426.	.0052338
85+	Total	M	82	338830.	43377.	.0163895
85+	Total	W	263	1076875.	77735.	.0052107
85+	Total	S	11	45489.	14069.	.0956645
85+	Total	NM	16	70079.	20564.	.0861066
85+	Total	Total	372	1531272.	101393.	.0043844
			SSI only			
Total	Total	Total	335	1550062.	125430.	.0065479
Total	Total	NM	123	546880.	62646.	.0131221
Total	Total	S	80	397264.	43744.	.0121251
Total	Total	W	61	249210.	42864.	.0295840
Total	Total	M	71	356709.	45562.	.0163148
< 18	Male	NM	2	7361.	5246.	.5079297
< 18	Female	NM	1	4370.	4370.	1.0000000
< 18	Total	NM	3	11731.	6828.	.3387586
18-24	Male	NM	13	67973.	20382.	.0899115
18-24	Female	S	1	4271.	4271.	1.0000000
18-24	Female	NM	14	68475.	21556.	.0991042
18-24	Female	Total	15	72745.	21975.	.0912562
18-24	Total	NM	16	136448.	31575.	.0535512
18-24	Total	Total	28	140718.	31863.	.0512712
25-34	Male	M	4	17112.	8626.	.2541071
25-34	Male	NM	9	56268.	19663.	.1221226
25-34	Male	Total	13	73380.	19990.	.0742110
25-34	Female	M	7	30357.	13351.	.1934159
25-34	Female	W	1	2801.	2801.	1.0000000
25-34	Female	S	13	65411.	22161.	.1147831
25-34	Female	NM	21	101224.	24471.	.0584425
25-34	Female	Total	42	199792.	32211.	.0259932
25-34	Total	M	11	47468.	17949.	.1429876
25-34	Total	NM	30	157492.	30640.	.0378502
25-34	Total	Total	55	273171.	36880.	.0182268
35-44	Male	M	2	9521.	6759.	.5040373
35-44	Male	W	1	4726.	4726.	1.0000000
35-44	Male	S	4	20770.	10631.	.2619952
35-44	Male	NM	6	39912.	17092.	.1833900
35-44	Male	Total	13	74928.	23953.	.1021975
35-44	Female	M	7	35734.	13694.	.1468636
35-44	Female	S	18	83043.	21535.	.0672484
35-44	Female	NM	7	32351.	12341.	.1455173
35-44	Female	Total	32	151128.	30387.	.0404283
35-44	Total	M	9	45255.	18444.	.1661076
35-44	Total	S	18	103813.	24016.	.0535186
35-44	Total	NM	13	72262.	21081.	.0851084
35-44	Total	Total	45	226056.	38468.	.0289576
45-54	Male	M	6	27401.	11254.	.1686981
45-54	Male	NM	3	16536.	9654.	.3408730
45-54	Male	Total	9	43936.	14828.	.1138946
45-54	Female	M	11	45134.	12658.	.0786538
45-54	Female	W	5	22396.	10125.	.2044074
45-54	Female	S	17	78309.	16748.	.0457406
45-54	Female	NM	7	32688.	12423.	.1444498
45-54	Female	Total	40	178526.	28290.	.0251104
45-54	Total	M	17	72535.	16176.	.0497318
45-54	Total	NM	10	49223.	15733.	.1021679
45-54	Total	Total	49	222462.	31375.	.0198909
55-64	Male	M	6	27229.	11135.	.1672420
55-64	Male	S	4	22691.	11438.	.2541047
55-64	Male	NM	5	30260.	14131.	.2180876
55-64	Male	Total	15	80179.	20680.	.0665257
55-64	Female	M	9	42124.	16624.	.1557495
55-64	Female	W	10	46112.	14711.	.1017860
55-64	Female	S	17	73164.	15898.	.0472172
55-64	Female	NM	1	5130.	5130.	1.0000000

Table I.—Variance estimates for OASDI and SSI participants under SSA-administered programs—Continued

Age	Sex	Marital status[1]	Unweighted count	Estimate	Standard error	Relative variance
			SSI only—cont.			
55-64	Female	Total	37	166529.	26792.	.0258833
55-64	Total	M	15	69353.	20930.	.0910775
55-64	Total	S	21	95855.	19367.	.0408220
55-64	Total	NM	6	35389.	15033.	.1804569
55-64	Total	Total	52	246708.	35316.	.0204914
65-69	Male	M	6	27450.	13480.	.2411725
65-69	Male	S	1	5738.	5738.	1.0000000
65-69	Male	NM	3	10665.	6212.	.3393365
65-69	Male	Total	10	43852.	15913.	.1316878
65-69	Female	M	6	25670.	10548.	.1688572
65-69	Female	W	10	39949.	13637.	.1165299
65-69	Female	S	4	18963.	9836.	.2690720
65-69	Female	NM	5	19067.	8551.	.2011198
65-69	Female	Total	25	103648.	20832.	.0403968
65-69	Total	M	12	53120.	20067.	.1427083
65-69	Total	S	5	24701.	11388.	.2125446
65-69	Total	NM	8	29731.	10569.	.1263746
65-69	Total	Total	35	147500.	28171.	.0364758
70-74	Male	M	7	26507.	10149.	.1465923
70-74	Male	NM	2	10523.	7442.	.5002612
70-74	Male	Total	9	37030.	12585.	.1155128
70-74	Female	M	3	12172.	7083.	.3386633
70-74	Female	W	6	24366.	9978.	.1677108
70-74	Female	S	3	16302.	9415.	.3335978
70-74	Female	NM	3	12947.	7512.	.3366193
70-74	Female	Total	15	65786.	18699.	.0807925
70-74	Total	M	10	38679.	15046.	.1513221
70-74	Total	NM	5	23470.	10574.	.2030004
70-74	Total	Total	24	102816.	25600.	.0619948
75-84	Male	M	5	19544.	8793.	.2024056
75-84	Male	W	3	8736.	5046.	.3336572
75-84	Male	Total	8	28280.	10138.	.1285093
75-84	Female	M	2	7917.	5598.	.5000312
75-84	Female	W	17	71632.	17733.	.0612834
75-84	Female	S	1	3901.	3901.	1.0000000
75-84	Female	NM	4	23433.	19539.	.6952958
75-84	Female	Total	24	106883.	27254.	.0650218
75-84	Total	M	7	27461.	13089.	.2271973
75-84	Total	W	20	80368.	19766.	.0604910
75-84	Total	Total	32	135163.	33839.	.0626804
85+	Male	S	1	4704.	4704.	1.0000000
85+	Female	M	1	2840.	2840.	1.0000000
85+	Female	W	8	28493.	11111.	.1520652
85+	Female	NM	2	7703.	5467.	.5038455
85+	Female	Total	11	39036.	12705.	.1059296
85+	Total	Total	12	43740.	13548.	.0959363
			OASDI and SSI			
Total	Total	Total	366	1598359.	152132.	.0090592
Total	Total	NM	51	243820.	33439.	.0188086
Total	Total	S	59	259581.	37829.	.0212375
Total	Total	W	168	701867.	69525.	.0098125
Total	Total	M	88	393092.	74110.	.0355438
18-24	Male	NM	2	8441.	5993.	.5040591
18-24	Female	NM	4	18518.	9315.	.2530180
18-24	Total	NM	6	26959.	11076.	.1687959
25-34	Male	S	1	10068.	10068.	1.0000000
25-34	Male	NM	7	33532.	10389.	.0959927
25-34	Male	Total	8	43600.	14467.	.1100987
25-34	Female	W	1	3580.	3580.	1.0000000
25-34	Female	NM	4	17978.	8990.	.2500436
25-34	Female	Total	5	21557.	9676.	.2014712
25-34	Total	NM	11	51510.	13738.	.0711380
25-34	Total	Total	13	65157.	17404.	.0713514
35-44	Male	NM	4	20395.	10223.	.2512503
35-44	Female	W	1	4870.	4870.	1.0000000

Table I.—Variance estimates for OASDI and SSI participants under SSA-administered programs—Continued

Age	Sex	Marital status[1]	Unweighted count	Estimate	Standard error	Relative variance
			OASDI and SSI—cont.			
35-44	Female	S	3	11948.	6915.	.3349714
35-44	Female	NM	1	5543.	5543.	1.0000000
35-44	Female	Total	5	22360.	10112.	.2045137
35-44	Total	NM	5	25938.	11629.	.2010072
35-44	Total	Total	9	42755.	14379.	.1131078
45-54	Male	M	1	6263.	6263.	1.0000000
45-54	Male	W	1	4059.	4059.	1.0000000
45-54	Male	S	1	5157.	5157.	1.0000000
45-54	Male	NM	5	25960.	13638.	.2759768
45-54	Male	Total	8	41439.	16379.	.1562345
45-54	Female	M	1	3789.	3789.	1.0000000
45-54	Female	W	1	4022.	4022.	1.0000000
45-54	Female	S	6	31886.	13127.	.1694950
45-54	Female	NM	2	8454.	5995.	.5028209
45-54	Female	Total	10	48150.	15453.	.1029962
45-54	Total	M	2	10052.	7320.	.5302876
45-54	Total	W	2	8080.	5713.	.5000105
45-54	Total	S	11	37043.	11909.	.1033558
45-54	Total	NM	10	34414.	14897.	.1873841
45-54	Total	Total	18	89589.	22334.	.0621503
55-64	Male	M	6	25913.	12198.	.2215964
55-64	Male	W	1	4987.	4987.	1.0000000
55-64	Male	S	2	10625.	7717.	.5276068
55-64	Male	NM	3	15120.	8737.	.3339024
55-64	Male	Total	12	56643.	17594.	.0964783
55-64	Female	M	8	38486.	14040.	.1330844
55-64	Female	W	11	46099.	14788.	.1029058
55-64	Female	S	9	34385.	12596.	.1341939
55-64	Female	NM	2	9177.	6489.	.5000067
55-64	Female	Total	30	128146.	23980.	.0350169
55-64	Total	M	14	64399.	20216.	.0985467
55-64	Total	W	12	51085.	16923.	.1097383
55-64	Total	S	12	45010.	14772.	.1077163
55-64	Total	NM	5	24296.	10883.	.2006359
55-64	Total	Total	42	184789.	32842.	.0315871
65-69	Male	M	12	53931.	17970.	.1110225
65-69	Male	W	2	7523.	5437.	.5222957
65-69	Male	S	1	6603.	6603.	1.0000000
65-69	Male	Total	15	68057.	18906.	.0771726
65-69	Female	M	6	24831.	8618.	.1204502
65-69	Female	W	32	129568.	26794.	.0427633
65-69	Female	S	5	22668.	10161.	.2009360
65-69	Female	NM	3	12794.	7440.	.3382045
65-69	Female	Total	46	189861.	29768.	.0245832
65-69	Total	M	18	78762.	22078.	.0785764
65-69	Total	W	34	137091.	29934.	.0476783
65-69	Total	S	6	29271.	12118.	.1713932
65-69	Total	Total	61	257917.	37955.	.0216558
70-74	Male	M	8	31406.	10147.	.1043939
70-74	Male	W	3	11621.	6777.	.3401275
70-74	Male	S	2	8966.	6391.	.5080770
70-74	Male	NM	3	15018.	8770.	.3410458
70-74	Male	Total	16	67010.	20146.	.0903885
70-74	Female	M	11	50253.	17738.	.1245843
70-74	Female	W	39	163619.	29621.	.0327747
70-74	Female	S	13	54596.	15206.	.0775686
70-74	Female	NM	4	16552.	8410.	.2581784
70-74	Female	Total	67	285020.	43907.	.0237310
70-74	Total	M	19	81659.	21322.	.0681784
70-74	Total	W	42	175240.	31815.	.0329614
70-74	Total	S	15	63562.	16201.	.0649655
70-74	Total	NM	7	31570.	12151.	.1481469
70-74	Total	Total	83	352029.	51120.	.0210879
75-84	Male	M	19	83750.	27374.	.1068347
75-84	Male	W	8	39519.	14358.	.1320007
75-84	Male	S	3	11340.	6703.	.3494227
75-84	Male	NM	1	4216.	4216.	1.0000000
75-84	Male	Total	31	138824.	30551.	.0484303

Table I.—Variance estimates for OASDI and SSI participants under SSA-administered programs—Continued

Age	Sex	Marital status[1]	Unweighted count	Estimate	Standard error	Relative variance
			OASDI and SSI—cont.			
75-84	Female	M	11	49022.	15289.	.0972771
75-84	Female	W	37	163484.	28646.	.0307027
75-84	Female	S	8	34864.	12522.	.1289945
75-84	Female	NM	6	24888.	8451.	.1153048
75-84	Female	Total	62	272257.	39936.	.0215161
75-84	Total	M	30	132771.	39096.	.0867091
75-84	Total	W	45	203003.	32494.	.0256211
75-84	Total	S	11	46204.	14203.	.0944941
75-84	Total	NM	7	29103.	9444.	.1053015
75-84	Total	Total	93	411081.	58833.	.0204830
85+	Male	M	3	15476.	5219.	.1137502
85+	Male	W	5	22409.	10090.	.2027545
85+	Male	S	1	6166.	6166.	1.0000000
85+	Male	Total	9	44050.	12925.	.0861010
85+	Female	M	2	9975.	7060.	.5010565
85+	Female	W	26	96512.	17763.	.0338759
85+	Female	S	2	10312.	7744.	.5639426
85+	Female	NM	2	7238.	5122.	.5007861
85+	Female	Total	32	124036.	22002.	.0314648
85+	Total	M	5	25450.	8883.	.1218274
85+	Total	W	31	118920.	20795.	.0305792
85+	Total	S	3	16477.	9898.	.3608795
85+	Total	Total	41	168085.	26407.	.0246826

[1]NM = Never married; S = Separated; W = Widowed; M = Married.

Chart I.—Half-sample assignment for pseudo-unit 1 cases

Stratum	Half-sample
1	11111110111010011011100011010110100011101001010011100010011010001000000
2	11111101110100110111000110101101000111010010100111000100110100010000001
3	11111011101001101110001101011010001110100101001110001001101000100000011
4	11110111010011011100011010110100011101001010011100010011010001000000111
5	11101110100110111000110101101000111010010100111000100110100010000001111
6	11011101001101110001101011010001110100101001110001001101000100000011111
7	10111010011011100011010110100011101001010011100010011010001000000111111
8	01110100110111000110101101000111010010100111000100110100010000001111111
9	11101001101110001101011010001110100101001110001001101000100000011111110
10	11010011011100011010110100011101001010011100010011010001000000111111101
11	10100110111000110101101000111010010100111000100110100010000001111111011
12	01001101110001101011010001110100101001110001001101000100000011111110111
.	
.	
.	
66	00000011111110111010011011100011010110100011101001010011100010011010001
67	00000111111101110100110111000110101101000111010010100111000100110100010
68	00001111111011101001101110001101011010001110100101001110001001101000100
69	00011111110111010011011100011010110100011101001010011100010011010001000
70	00111111101110100110111000110101101000111010010100111000100110100010000
71	01111111011101001101110001101011010001110100101001110001001101000100000

NOTE 2: Evaluation of Direct Variance Estimates from the 1984 SIPP Public Use File

Case weights and variable values were based on the rotation group as shown below:

Rotation group	Month(*)
1	4
2	3
3	2
4	1

The variables are referred to by their public use file names. (Starting character position of the month-1 field is shown in parentheses.)

1. Age 16 and over
   AGE-*(2206)>16.

2. Low Income Cash Only (LICO)
   H*TOTINC(178)<H*POV$(173).

3. LICO plus government noncash transfers (LICNC)
   H*TOTINC(178)+H*NONCSH(215)<H*POV(173).

4. Receiving Unemployment Compensation (UNCO)
   IO5AMT*(3820)+IO6AMT*(3848)+IO7AMT*(3876)>0.

5. Receiving Cash from a means tested program (CBPR)
   H*-TRAN(201)>0.

6. Receiving food stamps (FS)
   H*-FDSTP(251)>0.

7. Receiving noncash benefits other than food stamps (NCBPR)
   CAIDCOV*(2672)=1, or
   H*PUBAMT(258)>0, or
   H*-LUNCH(266)≠0, or
   H*-BREAK(267)≠, or
   H*-4804(269)>0, or
   H*NONSCH(215)>H*-FDSTP(251).

8. Some labor force activity (SLFA)
   ESR-*(2593)≥1, and
   ESR-*(2593)≤7.

9. Hispanic (HIS)
   ETHNICTY(2278)≥14, and
   ETHNICTY(2278)≤20.

Evaluation of Direct Variance Estimates from the 1984 SIPP Public Use File

INTRODUCTION

The 1984 public use data files of the Survey of Income and Program Participation (SIPP) provide pseudo stratum and pseudo primary sampling unit codes that permit direct estimates of sampling variances by a number of methods. The actual sample design parameters are withheld from public use to prevent access to small geographic areas where disclosure of individual identities might be possible. The Social Security Administration (SSA) has used the pseudo codes to compute sampling variances for SSA program participants. (Bye and Gallicchio 1988.) Although the variance estimates appeared to be reasonably well behaved, no external assessment of them was made.

In this note we report the results of a comparison of direct variance estimates from the public use file with variance estimates based on the original sample design computed by the Bureau of the Census. The comparison involves estimates of 36 population totals that comprised the item set for the first generalized curve ("program participation and benefits, poverty") in the SIPP User's Guide (1987, page 7-5). The SSA direct variance estimates were computed using 72 balanced half samples derived from the pseudo design. Details are provided in Bye and Gallicchio (1988). The Census estimates were obtained from a set of 50 half samples that were not fully balanced derived from the original design. Case weights in each of the Census half samples were adjusted to a common set of population totals, replicating the weighting methodology of the full sample. The SSA half sample case weights were constructed by multiplying the full sample weight by 2.

The results of the comparison are very encouraging. Most of the items compared showed small differences in coefficient of variation (CV). The differences were both positive and negative with no apparent pattern. This finding together with the ease of computation of the estimator makes the direct estimation of variances from the public use sample very attractive to the data analyst.

VARIANCE ITEMS

This section presents the SSA item specifications. (An exact match of public use file estimates with those provided by Census was not expected because the Census estimates were produced some years ago from an internal file for which specifications are not longer available.) The 36 items were combinations of 9 characteristics (Bureau of the Census, 1985). SSA's construction of these characteristics relate to individual and household status as of September 1983.

RESULTS

Table 1 presents the comparison of Census and SSA variance estimates for the 36 items. As expected the estimated totals do not agree exactly, and these differences contribute to the differences in estimated standard errors. A more meaningful comparison, therefore, is the ratio of CVs. With the exception of items 26 and 32, the Census and SSA CVs are quite similar. The ratios of the SSA CV to Census CV range from a low of .849 to a high of 1.093. There is no apparent pattern to the differences as a function of size of the estimate.

The SSA CV for item 26 (item 32 consists of essentially the same sample cases as 26) is about 50 percent larger than the corresponding Census CV. An examination of the 72 SSA half sample estimates of this characteristics (data not shown here) indicates a wide range of estimated totals but no extreme outliers. The size of the CV for this estimate appears to be a chance occurrence indicating, perhaps, that the SSA variance estimator might have a larger variance than the Census estimator, especially when cells are small. A comparison of substantially more items would be needed to investigate this further.

REFERENCES

Bye, Barry V. and Gallicchio, Salvator J., "A Note on Sampling Variance Estimates for Social Security Program Participants From the Survey of Income and Program Participation," Social Security Bulletin, Vol. 51 No. 10, October 1988.

Survey of Income and Program Participation, User's Guide, Bureau of the Census, Department of Commerce, July 1987.

Memorandum for Documentation from Karen E. King, Subject: SIPP Variances: Items for Generalized Variance Parameters, Bureau of the Census, Department of Commerce, June 15, 1985.

Table 1. Census/SSA Variance Comparison

	Census			SSA			Ratio SSA/Census		
	Estimate	St. Err.	CV	Estimate	St. Err.	CV	Estimate	St. Err.	CV
1. low income cash (lico) only	20875350	693897	0.033	21459181	707972	0.033	1.028	1.020	0.993
2. lico plus govt noncash trans (licnc)	19791310	684505	0.035	20373140	701869	0.034	1.029	1.025	0.996
3. receiving unemp. compensation (unco)	3038883	154221	0.051	2916786	156861	0.054	0.960	1.017	1.060
4. receiving cash from mean-test (cbpr)	14556060	703225	0.048	14546133	651057	0.045	0.999	0.926	0.926
5. receiving food stamps (fs)	11920830	551678	0.046	12006268	535354	0.045	1.007	0.970	0.964
6. receiving noncash o/t fs (ncbpr)	25730990	875913	0.034	24921228	898371	0.036	0.969	1.026	1.059
7. some labor force activity (slfa)	112328300	1884640	0.017	115087192	1901174	0.017	1.025	1.009	0.985
8. slfa in hh with licnc	9497841	324529	0.034	9625030	354010	0.037	1.013	1.091	1.076
9.slfa receiving unco	2913951	149957	0.051	2799895	155409	0.056	0.961	1.036	1.079
10.slfa in hh receiving fs	5516297	258996	0.047	5469103	246196	0.045	0.991	0.951	0.959
11.no labor force act (nlfa) in hh lico	10841920	445140	0.041	11296831	455121	0.040	1.042	1.022	0.981
12.nlfa receiving unco	124932	24073	0.193	116890	23229	0.199	0.936	0.965	1.031
13.nlfa in hh receiving cbpr	8674453	432411	0.050	8810027	434336	0.049	1.016	1.004	0.989
14.same as 3. in hh with lico	519714	66145	0.127	556148	64921	0.117	1.070	0.981	0.917
15.same as 4. in hh with lico	6949680	450128	0.065	6899747	412336	0.060	0.993	0.916	0.923
16.same as 5. in hh with lico	8561362	454857	0.053	8514647	431563	0.051	0.995	0.949	0.954
17.same as 6. in hh with lico	11244680	507620	0.045	10558880	500328	0.047	0.939	0.986	1.050
18.same as 7. in hh with lico	10033130	323310	0.032	10162351	358123	0.035	1.013	1.108	1.094
19.same as 9. in hh with lico	502881	63081	0.125	539965	64644	0.120	1.074	1.025	0.954
20.same as 10. in hh with lico	3628493	208210	0.057	3485122	187039	0.054	0.960	0.898	0.935
21.same as 12. in hh with lico	16833	8519	0.506	16183	8105	0.501	0.961	0.951	0.990
22.same as 13. in hh with lico	4588350	306016	0.067	4686842	310529	0.066	1.021	1.015	0.993
23. number of hispanics 16+	9606953	625095	0.065	9461055	651284	0.069	0.985	1.042	1.058
24. his low income cash (lico) only	2371371	274453	0.116	2247679	275540	0.123	0.948	1.004	1.059
25.his lico plus govt noncash trans (licnc)	2212253	258185	0.117	2112670	262870	0.124	0.955	1.018	1.066
26. his receiving unemp. compensation (unco)	235565	35162	0.149	240340	52543	0.219	1.020	1.494	1.465
27. his receiving cash from mean-test (cbpr)	1704630	212180	0.124	1669796	203091	0.122	0.980	0.957	0.977
28. his receiving food stamps (fs)	1602333	210802	0.132	1555791	187125	0.120	0.971	0.888	0.914
29. his receiving noncash o/t fs (ncbpr)	3090985	316105	0.102	3116432	330621	0.106	1.008	1.046	1.037
30. his some labor force activity (slfa)	6231358	385368	0.062	6201126	408599	0.066	0.995	1.060	1.065
31. his slfa in hh with licnc	1008518	123187	0.122	988474	130064	0.132	0.980	1.056	1.077
32.his slfa receiving unco	218963	32774	0.150	227275	51143	0.225	1.038	1.560	1.503
33.his slfa in hh receiving fs	672542	109441	0.163	682383	94327	0.138	1.015	0.862	0.849
34. his no labor force act (nlfa) in hh lico	1296891	161290	0.124	1202075	154376	0.128	0.927	0.957	1.033
35.his nlfa receiving unco	16596	8850	0.533	13065	7613	0.583	0.787	0.860	1.093
36.his nlfa in hh receiving cbpr	1054751	126333	0.120	1000704	128222	0.128	0.949	1.015	1.070

Survey of Income and Program Participation

QUALITY OF SIPP ESTIMATES

8901 78

Rajendra Singh, Lynn Weidman, Gary M. Shapiro
U.S. Bureau of the Census

April 1989

A preliminary version of this paper was presented at the "International Symposium on Panel Surveys," November 19-22, 1986 in Washington, D.C. A revised and expanded paper was presented at the Social Science Research council's "Conference on Individuals and Families in Transition: Understanding Change through Longitudinal Data, March 16-18, 1988, in Annapolis, Maryland.

The authors wish to express their appreciation for the work of Kimberly Wilburn and Cora Wisniewski, without whose dedication and willing attitude this paper could not have been completed. Due to their exceptional efforts and skill; we were able to dispatch a lengthy study to meet an imposed deadline. We would also like to thank Mary Ellen Beach, Jack McNeil, Daniel Kasprzyk, Norman Johnson, and Randall Parmer, all of the Census Bureau, for their valuable comments. The views expressed are the authors' and do not necessarily reflect those of the Census Bureau.

TABLE OF CONTENTS

I. Introduction..... 1

A. Background..... 1
B. Sample Design..... 3

II. Quality of Estimates..... 4

A. Quality of Core Items..... 5
B. Quality of Estimates from Topical Modules..... 7
C. Quality of Gross Flows and Length of Spell Estimates..... 9

III. Error Sources..... 13

A. Identification of Sources..... 13
B. Nonresponse and Coverage of Population..... 16
C. Examination of Error Sources..... 18

IV. How Estimates Can Be Improved..... 20

A. Research for Improved Understanding..... 21
B. Research for Improving Estimates..... 28

V. Summary..... 34

TABLES
REFERENCES

QUALITY OF SIPP ESTIMATES

Rajendra Singh, Lynn Weidman, Gary M. Shapiro
U.S. Bureau of the Census

I. INTRODUCTION

The Bureau of the Census has been conducting interviews for the Survey of Income and Program Participation (SIPP) since October 1983. The SIPP is a national survey and is designed to provide improved information on income and participation in government programs for the noninstitutionalized United States population. Person and household characteristics that may influence income and program participation are also available from the SIPP. This information is vital for improving the capability of federal agencies to formulate and evaluate their policies and programs in the areas of income and social welfare.

A. Background

The estimates produced from the survey can be divided into two groups. The first group includes cross-sectional and cross-sectional type estimates. These estimates are obtained from the wave data files and the longitudinal data files. Examples of such estimates from wave files include the unemployment rate in March 1987, net change in unemployment rate between March 1986 and March 1987, number of persons participating in the Food Stamp Program in February 1987, and the number of females who completed high school in 1986. Annual estimates of income and estimates of change of certain characteristics are examples from longitudinal files. For our discussion, these estimates will be called cross-sectional estimates. The method developed for producing wave file estimates is described in King (1985), and King and Kim (1986). The estimation method developed for the first SIPP longitudinal file covering the first three interviews of the 1984 panel is presented in Kobilarcik and Singh (1986). The methods for the longitudinal 1984 panel file are presented in Hock (1988).

The second group includes the estimates of gross flows (transition from one state of economic or labor condition to another state) and distributions of the length of spells. The transition from any state, say 'A', to another state, say 'B' triggers an end to spell of state 'A' and the beginning of a spell for state 'B'. Thus, an estimate of gross flows has a direct effect on spell estimates. These estimates are important because they could serve as a very powerful instrument in explaining socio-economic processes. For example, what happens to the health insurance coverage of a person who no longer receives welfare benefits?

In this paper, we discuss quality issues for both cross-sectional and gross flow/spell estimates. We discuss what we know about the quality of the SIPP data, the different types

of error that can occur, and ideas for research to better understand and reduce error. A major purpose of the paper is to strongly encourage people outside the Census Bureau to research ideas discussed here and on other ideas that will improve our understanding of the quality of the estimates and help improve it.

We will first give a summary of the major points in the paper. We begin with a brief description of the SIPP sample design. Section II discusses in detail what we know about the quality of SIPP estimates. For cross-sectional core data, the SIPP estimates of the number of recipients for government programs and amounts of income received are generally lower than available independent estimates from administrative sources. However, SIPP estimates related to programs are generally closer to the independent estimates than are Current Population Survey (CPS) estimates. In particular, based on initial evaluation the SIPP estimates of persons below the poverty level may be superior to the CPS estimates.

Little information is available about estimates of change. There has been some evaluation of topical module data. A couple of apparent problems with this data have been uncovered. The apparent problems are 1) The educational financing data seems to be of generally poor quality; and 2) The characteristics of tax filers in SIPP are different from IRS data.

For gross flow and spell estimates from the core data there is one particular problem. Many more changes in recipiency status and amounts occur between a pair of two consecutive months in a different interview than between two months within the same interview. We have examined three income sources to see if the start-up and exit rates are biased by this problem. For food stamps, there is no evidence of bias. For aid to families with dependent children (AFDC), sampling errors are too large to be able to draw reliable conclusions. For supplemental security income (SSI), start-up and exit rates do appear to be significantly biased. Thus, the quality of these rates appear to vary by income source. For some purposes, eg. multivariate analysis at the micro level of gross flow and spell estimates, the affect of this inconsistency problem is unknown. More evaluation of micro level relationship among variables is needed to judge the quality of SIPP data for its uses in multivariate analysis.

Section III briefly discusses a number of different error sources. Some appear to have minor effects on estimates and some have at least the potential to cause major effects on some estimates. The sources of minor effect are interviewer coding, data coding, and use of proxy respondents. The potentially major effect sources are changes in interviewer, nonresponse, undercoverage, imputation, questionnaire wording and design, length of recall, and learning effects of respondents.

We continue section III by discussing three studies that have examined some of the sources of error. In a recall effect study (Petroni, 1986), we concluded that for many questions respondents tend to give the same response for all four months covered by a single interview. In a transition pattern study (Weidman, 1986 and 1987), we concluded that transitions did not seem to differ much among demographic groups and by self vs. proxy respondent. However, transitions are greater when some of the data has to be imputed. For the third study (McArthur and Short, 1986), we looked at the characteristics of people who remained as respondents and those who became noninterviews after responding in earlier interviews.

Section IV of the paper discusses a number of ideas for research. There are 12 research proposals aimed at improving our understanding of quality and 14 proposals for improving estimates themselves. Some examples of areas for research to improve understanding are: time-in-sample bias, expanding reinterviews, and coverage research. Some examples for improving estimates are: reducing complexity, reducing nonresponse, changing the reference period, increasing respondent effort, and improving interest and dividend incomes. Section V presents a brief summary of the paper.

Section II of the paper makes it clear that there are major gaps in our knowledge about the quality of the SIPP estimates. Even if we were to do all the research discussed in Section IV, we would only close some of the gaps. With the amount of data that can be provided from the SIPP and the disparity in the uses that can be made of it, it would be impossible to make a simple overall statement of the quality and adequacy of the estimates even if we knew everything possible about quality. It is also obvious that only a few of the research areas of section IV can be substantially addressed by Census Bureau staff in the short term. Although we hope that people outside the Bureau will address a few areas as well, this will still leave a lot of important research undone.

B. Sample Design

The SIPP is a multistage stratified systematic sample of the noninstitutionalized resident population of the United States. This population includes persons living in group quarters, such as dormitories, rooming houses, and religious group dwellings. Noncitizens of the United States who work or attend school in this country and their families are eligible. this country and their families are eligible. Crew members of merchant vessels, Armed Forces personnel living in military barracks, and institutionalized persons, such as correctional facility inmates and nursing home residents are ineligible. In addition to these general restrictions, only persons who were residing in the United States at the time of the first interview were eligible for SIPP. Also, only per-

sons who were at least 15 years of age were eligible for interview, although limited data on children were also collected by proxy interviews.

Initially, a sample of living quarters in selected Primary Sampling Units (PSUs) is taken. Living quarters are considered separate if the occupants do not live and eat with any other person in the structure and have either direct access from the outside of the building or through a common hall, or complete kitchen facilities for that unit only.

The SIPP sample is divided into four groups of equal size called rotation groups. One rotation group is interviewed each month. In general, one cycle of four rotation groups is called a wave. This design provides a smooth and steady work load for data collection and processing. Persons in the sample are interviewed once every four months for approximately two and one-half years. The reference period for the interview is the four months preceding the interview month. For example, for the first SIPP sample, the reference period for the November 1983 interview month (rotation group 2) was July through October 1983. These sample persons were interviewed again in March 1984 for the November 1983 through February 1984 period.

Persons 15 years old and over present as household members at the time of first interview are to be part of the survey for the entire two and one-half year period. With certain restrictions, these sample persons are followed if they move to a new address. "New" persons living with sample persons are considered to be part of the sample only while residing with these sample persons. More details on the SIPP design are given in Nelson, McMillen, and Kasprzyk (1985).

The SIPP questionnaire is long and complex. Questions are asked by specific type of cash and non-cash income on months received and amounts per month. For many types of income, additional questions are asked of recipients. For example, in households with children covered by medicaid, up to 8 questions about health insurance are asked. Questions are also asked about assets and labor force status. Topical modules on various subjects are also included in most interviews.

II. QUALITY OF ESTIMATES

The quality of the SIPP estimates is judged by comparing them with estimates from independent sources primarily to evaluate bias. These independent sources include administrative records maintained by various government agencies and household surveys conducted by government agencies and other survey organizations. The magnitude of nonsampling errors varies from source to source and makes it difficult to compare estimates. Furthermore, the estimates for the SIPP are produced only for the 1984 panel,

which may be different because it's the first one. Therefore, the results presented here should be considered preliminary and caution should be exercised in drawing conclusions about the quality of the SIPP estimates.

A. Quality of Core Items

Data on a large number of items are collected in each SIPP interview. These items are called core items and two different types of estimates - rates (or percents) and totals - are produced from them. Estimates of change are also produced for each of these. The quality of these estimates is discussed below.

1. Estimates of Rates and Levels

The quality of selected cross-sectional estimates based on the core part of the questionnaire is discussed in this section. The selected estimates primarily represent income and program participation items and include income from wage and salary, food stamps, social security, etc. Table 1 presents quarterly SIPP and 1983 CPS estimates as a percent of independently derived estimates. The table shows that, except for wage and salary income, estimates derived from SIPP are higher than the corresponding 1983 CPS estimates and are better than the CPS assuming that the independent estimates are accurate. However, these estimates are lower than those for the corresponding independent source, except for social security income.

A careful examination of these estimates also suggests that SIPP provides better estimates of number of program participants than it does of aggregate income amounts for 1) veteran's compensation or pension and 2) food stamps. These results suggest that either the income amounts for these two programs tend to be underreported by beneficiaries or the beneficiaries with larger amounts are disproportionaly underrepresented in the SIPP. The administrative record check study currently underway at the Census Bureau may shed light on this issue (Moore 1986). Furthermore, the quality of estimates other than unemployment compensation appears to be quite stable over time (see tables 1 and 2). Coder (1987b) monitored estimates of state unemployment compensation for all quarters through unemployment compensation for all quarters through 1985 from the SIPP 84 panel and found that their quality appears to be declining. These quarterly estimates are presented in table 2.

Carlson and Dalrymple (1986) compared selected income characteristics of food stamp recipients from two data sources: Wave 1 of the 1984 SIPP Panel and the Food and Nutrition Survey (FNS) of administrative records of food stamp participants in August 1983. Those who were identified as food stamp recipients in SIPP for September

1983 were analyzed in their study. (They felt this time difference should not adversely affect their study since their comparison between the SIPP August and September 1983 reference month files showed trivial differences.) They found that the differences in income characteristics between the SIPP and the FNS estimates were relatively small for the households with only one food stamp unit and no subunit. However, SIPP showed considerably fewer households (36%) with both Aid to Families with Dependent Children (AFDC) and food stamps than the FNS (46%).

When households with subunits were included in the analysis, they found larger differences for selected income characteristics. Some of the differences could be explained by the relative influence of characteristics of the members in subunits. However, the differences were not entirely explained.

The quality of the SIPP poverty rate was evaluated by comparing it with the CPS rate. Note that the concepts and the procedures for the CPS are different from the SIPP and the comparison of their estimates is not totally valid. Coder et. al. (1987) obtained the CPS type income estimates for the SIPP in order to compare SIPP with CPS. Annual SIPP household income was determined using the household composition as it was for the twelfth reference month on the longitudinal research file consisting of the first three interviews in the 1984 SIPP panel. He showed that SIPP estimates lower poverty rates than CPS for all persons, white and black. The poverty rates for all persons from the SIPP and the CPS were 13.0% and 14.8%, respectively. The rates for white and black also showed similar differences. Ruggles and Williams (1986) also found lower poverty rates by family type for the SIPP than the CPS using the CPS type income estimates and the SIPP data for waves 2 through 5 from the 1984 SIPP panel. We believe the poverty rates from the SIPP may be better since SIPP captures income from transfer programs better than CPS (see table 1). SIPP is also more successful in capturing persons with marginal income because of a shorter recall period.

Vaughan (1988) compared interest and divided income amounts from the SIPP with the CPS and independent estimates. The SIPP provided better dividend amount data than the CPS. However, the estimates from both surveys were way too low compared to the independent estimates. The SIPP and the CPS both underestimated income amounts from interest. Data did not show which of the two was better.

Evaluation of the estimates produced from the longitudinal data file is in its early stages. Tables 3-6 present a few selected estimates from Coder (1986b). These estimates have been compared with estimates from independent

sources. Some estimates appear to be of good quality - for example, persons receiving AFDC, food stamps in fourth quarter of 1983, mean annual income amounts from rents and royalties - although more research is needed.

2. Quality of Estimates for Change

SIPP also provides estimates of change in level (or percent) for many characteristics, such as the number of food stamp participants and the number of households by source of income. As a part of the SIPP evaluation, estimates of changes between the third quarters of 1983 and 1984 were examined for certain characteristics. Table 7 presents relative change estimates from the SIPP and independent sources. Differences in these estimates are also presented in the table. These differences between estimates from the SIPP and independent sources for Social Security, SSI, AFDC, and food stamps appear to be large for analytical purposes but they are not statistically different due to small SIPP sample size. (The changes in level estimates were also not statistically different.) However, the numbers of total households with four (out of five) selected assets are significantly lower for the third quarter of 1984 than for that of 1983. (see Table 8.) Further analysis utilizing either estimates for a longer period or estimates from independent sources will shed light on whether or not the change estimates are influenced by nonsampling errors such as time-in-sample bias, learning effects, etc.

Hill (1987) studied marital status and its changes over time as reported for the SIPP and independent data sources. Independent national estimates were based on either pertinent information in the Statistical Abstract (1986), a combination of published vital statistics and the CPS, or obtained from the Panel Survey of Income Dynamics (PSID). SIPP estimates were based on waves 1 through 3 data of the 1984 panel for rotation groups 1 through 3, individuals aged 15 and over responding in all three waves. Wave 3 weights were used since longitudinal weights were not available. Hill found significantly lower proportions of changes in marital status reported in SIPP over the course of the year than for the other sources. For example, for persons 15 years or older, SIPP reported 1.4% becoming married, while the Statistical Abstract (1986) indicated 2.6% becoming married. SIPP reported 0.6% becoming divorced while a combination of Vital Statistics and the CPS reported 1.3%. Lower changes were reported for all status changes except into widowhood.

B. Quality of Estimates from Topical Modules

SIPP is designed to provide data on a number of special topics. The data on these special topics (usually called

topical modules) are not collected during each interview. The evaluation of the topical module data is not completed and it would be difficult to discuss here the quality of data from each topical module evaluated so far. However, the quality of data for selected modules will be discussed. Since the quality cussed. Since the quality of the data from a topical module depends on its topic, no general conclusions about the quality of topical module data is possible at this time.

SIPP collected data in Wave 5 of the 1984 panel on child care arrangements. The data analyzed were averages of the usual child care arrangements from December 1984 through March 1985 and the results were presented in the Current Population Reports, Series P-70, No. 9, of the Census Bureau. The report also compared the SIPP data with May 1985 data from the CPS and 1984 individual income tax returns. A few of these comparisons are presented here. SIPP estimates about 900,000 children under 15 years of age were cared for by unmarried men while CPS estimates that 671,000 children under age 12 and 528,000 children 12 to 17 years old were with unmarried fathers. Assuming a uniform distribution for children 12 to 17 years old that were cared for, the CPS estimate for children under 15 years of age that were cared for is 935,000. Thus, SIPP and CPS estimates appear to be comparable. SIPP and CPS estimate that 5.5% and 4.6%, respectively, of working women were absent from work due to failure in childcare arrangements.

SIPP estimates of employed women with at least 1 child under 15 and of child care arrangements don't seem to be that inconsistent with IRS estimates. (See Current Population Reports, Series P-70, No. 9.) However, inconsistencies between SIPP and IRS universes preclude any definite conclusions.

During Wave 4 of the 1984 SIPP panel, data on household wealth and asset ownership were collected. A comparison of the SIPP aggregate asset amounts with estimates derived from the Flow of Fund data of the Federal Reserve Board (FRB) along with the detailed analysis of the SIPP data is presented in the Current Population Reports, Series P-70, No.- 7 of the Census Bureau. Curtin et al (1987) compared the SIPP wealth data with the 1983 Survey of Consumer Finances (SCF) and the 1984 Wealth Supplement to the Panel Study of Income Dynamics (PSID). One should be cautious in interpreting their results. This is due to the fact that the SIPP data file has wealth top-coded. In addition, there are some conceptual and logical differences among these surveys.

Table 9 presents the estimates from the SIPP and the FRB data published in the Current Population report. The differences in estimates from the two sources are large, but one should be careful in drawing conclusions from this table due to the following limitations. 1) The household sector in FRB data include nonprofit organizations and private trusts not

covered under the SIPP. 2) The SIPP universe consists of noninstitutionalized resident population living in the United States and at least 15 years of age. The FRB Balance Sheet includes the asset holdings of the institutionalized population. 3) The household sector of the FRB balance sheet is estimated as a residual after allocations are made to farm business, nonfarm noncorporate business, nonfinancial corporate business and private financial institutions. As a result, accuracy of household sector estimates is reduced.

The Annual Roundup topical module was administered in Wave 6 of the 1984 panel. Coder (1987d) found that the SIPP estimate of 111.9 million recipients of wage and salary for calendar year 1984 is lower than the CPS estimate of 114.4 million (the SIPP and the CPS estimates include imputed data.) Furthermore, the overall nonresponse rate (including household, person and item nonresponse) for wage and salary amounts was about 40 percent. This rate is much higher than the CPS rate of 24%. Also, only 30% of the amounts were taken from W-2 forms even though its use was encouraged in SIPP. The data from the remaining respondents were based strictly on their recall. Table 10 presents median wage and salary income of those who used W-2 forms and those who did not. The table shows that, in general, the median income of those who used W-2 forms is higher than those who did not. Furthermore, SIPP estimates of wage and salary based on the core data are lower than the CPS estimate. (See table 1). Overall the quality of wage and salary data from the SIPP is not as good as from the CPS.

Coder (1987e) also found that the distribution of tax returns by return type in the SIPP is different from the IRS. He indicated that the number of single returns are underreported in the SIPP. Also the SIPP adjusted gross income (AGI) medians by return type are higher than for the IRS.

Kominski (1987) analyzed the data for educational financing collected in Wave 6 topical module of the 1984 panel and found that the estimates in general do not come close to independent estimates of financing for the period these data reference. (The topical module data he used was not edited.) He also observed large discrepancies in reporting the same phenomenon in the core and the topical module. Thus, the overall quality of the SIPP data for educational financing is poor in the 1984 Panel. Starting with the 1985 Panel, the questions related to educational financing were changed substantially so that the core questions closely mirror topical module questions.

C. Quality of Gross Flow and Length of Spell Estimates

Let us first discuss the measurement of gross flows between any pair of consecutive months. For example, in table 11, gross flows between January 1984 and February 1984 are

observed from a single interview (i.e., second interview) for rotations 2, 3, and 4. For rotation 1, they are observed by linking two interviews (the second and third interviews). Thus, the SIPP design produces four measurements, one for each rotation group. Three of them come from a single interview (within reference period) and one measurement comes from a pair of consecutive interviews.

The preliminary analysis of unweighted data from the SIPP [Coder 1986a] presents evidence that gross flows differ for pairs reported by the same interview from those reported from two consecutive interviews. Some selected results are presented in table 12, which shows month-to-month changes in recipiency and amounts for food stamps. Month-to-month changes for fourth to fifth and eight to ninth correspond to the seams where reference periods join (i.e., two consecutive interviews). All other pairs are from the same interview. Note that there are many more transitions between the eighth and ninth months and the fourth and fifth months than between other pairs of months. This pattern also holds for other characteristics such as railroad retirement, child support payments, state unemployment compensation, etc. [Coder 1986a]. Moore and Kasprzyk (1984) also observed similar results in ISDP-79 data for these and other characteristics. These differences are clearly due to nonsampling error in reporting. This reporting pattern affects estimates of the covariance structure and has significant adverse effects on multivariate analyses dealing with transitions or length of spells.

The problem with gross flow estimates is not unique to SIPP. Hill (1987b) also reported problems with gross flow estimates in the Panel Survey of Income Dynamic (PSID). Similar problems for the Current Population Survey have been known to analysts for over twenty years and are discussed in the proceedings of the Conference on Gross Flows in Labor Force Statistics (1985).

A large proportion of the research on transitions at the Census Bureau has concentrated on government benefit programs and labor force status. This work includes comparisons of SIPP with CPS and administrative data in order to evaluate the quality of reported transition rates, and examination of relationships between demographic characteristics and the months in which transitions are reported. In this section we review the results of the comparison studies.

Ryscavage and Feldman-Harkins (1988) compared gross flow and stock (level) estimates for labor force status from the SIPP and the CPS. In their study they found that the SIPP provided lower gross flow estimates than the CPS. The study found that the gross flow estimates from the SIPP were more consistent with the corresponding estimates of stocks (levels). They pointed out that this is bound to be the case because of the SIPP design. The larger inconsistency in the

CPS estimates was attributed to the fact that the gross flow estimates from the CPS for any pair of two consecutive months are obtained from two different interviews. They reserved their judgement about the quality of the SIPP labor force flows at this point since the survey designs in the SIPP and the CPS are very different and suggested further investigation before reaching any judgement.

Burkhead and Coder (1985), and Coder (1986a) show that transitions are dramatically understated most months and/or overstated every fourth month. If transitions are overestimated at the seams and underestimated within reference periods, then the combination of these for a given pair of months or over an interval of months may be less biased. With this in mind, studies to evaluate the bias in reporting for participation for food stamps (Judkins 1986), AFDC (Maher 1987b) and supplemental security income (SSI) (Maher 1987c) have been completed. In these studies, start up and exit rates (transition rates) for SIPP were computed using unweighted data from the SIPP longitudinal file (Coder 1986a). Noninterviewed cases were excluded and imputed data were used for item nonresponse.

Food stamp start-up and exit rates were computed from administrative record data prepared by the Urban Institute (1985) for the Food and Nutrition Service. These data were obtained using a two-stage stratified sample (with equal probability of selection) of local food-stamp offices in the 48 coterminous states and the District of Columbia. Complete case histories on subsamples of cases active between October 1, 1980 and December 31, 1983 were collected. Data from the last six months were used in the comparison study. Due to internal inconsistencies, about eight percent of the cases from the administrative records were discarded.

The start-up rate is defined to be the percent of active participants who are in the first month of a participation spell. Similarly, the exit rate is defined as the percent of active participants who are in the last month of a participation spell. The average rates were compared for four pairs of reference months for SIPP with six pairs of reference months (covering the same calendar months) for the administrative records. These results are presented in table 13. This study, even with its limitations, was very encouraging. Transition rates based on measures for all four rotation groups provide no evidence of differences between SIPP and Administrative Records Data. The results may be different if weighted data are used, but it seems unlikely.

For AFDC, estimates of the administrative record rates were obtained from several issues of Quarterly Public Assistance Statistics (1983,1984) which present data from complete sets of administrative records. Comparisons of average start-up and exit rates were made for the periods July-December 1983, October 1983-June 1984 and July 1983-June 1984 (see table

14). The average start-up rates are slightly lower for SIPP and the average exit rates are 20-30% lower for SIPP. When these differences are tested, they are not significant at the 10% level. (The tests were performed as if the estimates from each of the three periods are independent, but they have considerable overlap in data.) Standard errors on the SIPP estimates are very large, so no conclusions on the accuracy of transition rates are really possible. It is desirable to examine these estimates over a longer period of time in order to assess the bias in them.

For SSI, issues of the Social Security Bulletin (1984,1985) provide estimates of start-up rates for complete sets of administrative records, including people who are institutionalized or under age fifteen. Since SIPP does not include receipt of benefits for these people, adjustments to estimates from the bulletin were made based on the Social Security Administration's December 1983 1% file. Comparisons of average start-up rates are made for periods similar to those used in the AFDC study, and they indicate problems with the SIPP estimates (see table 15). Most of the within reference period rates for SIPP are as high or higher than all of the administrative rates, and the rates at the seams are still several times higher than those within waves. This results in tests that show significant differences at the 10% level between the two sources.

This higher start-up rate reported in SIPP could be a result of some confusion on the part of interviewed recipients between regular social security and SSI. If this is the case, then a comparison of exit rates should show the same pattern of monthly over-reporting as for start-up rates.

The results from these 3 studies suggest that each benefit source should be individually evaluated before using longitudinal estimates of transitions from SIPP. Similar types of studies should be extended to receipt and amount of income from various assets, as they show the same kind of within reference period vs. seam reporting pattern (Coder 1986a).

The reporting of more changes at the seam could have adverse effects on covariance structures and hence on micro-level analysis. The study of Young (1989) sheds some light on transition correlations between a number of different events and amount change status. Table 24 presents some of the correlations he computed. The number 1 in column 2 of the table refers to the pair of seam months, and numbers 2, 3, and 4 refer to the other 3 pairs formed by reference months within the interview. The correlations corresponding to these pairs are presented in their respective rows. Except for correlations of 'marital status' and 'married spouse present' with other characteristics, they did not show a pattern of distortion in bivariate relationships. These results are very encouraging. However, until more analysis is completed we should be careful reaching a definite conclusion.

Let us optimistically assume that other evaluation studies yield results similar to those for food stamps. Does it mean that our gross flows and length of spell estimates can be used by policy makers and social scientists? It depends on their goals. For some purposes they will be useful while for others they will not. For example, estimates of transitions based on measures for all four rotation groups for a given month at the macro level will be satisfactory. Furthermore, the estimate of change in number (or rate) of transitions and in length of spells based on measures for all four rotation groups would also be satisfactory if time-in-sample effect is small (compared to estimates). Such estimates would be worthwhile for policy makers and could assist them in evaluating their policies. On the other hand, more evaluation of covariance structures is needed to judge the usefulness of micro level multivariate analysis whose goal is to understand economic processes.

At present, very little is known about the bias in SIPP estimates. We need extensive research in this area to understand the problem better. Some possible research areas for determining the causes of the problem and how to correct it are discussed in Section IV.

III. ERROR SOURCES

A. Identification of Sources

In order to conduct research into alleviating the problems discussed previously, we first attempt to identify causes for the observed response patterns. These causes can be separated into two types: those related to the respondents and those related to the survey instrument and its processing. Of course, there is some overlap between these types. The latter type includes questionnaire wording/design, interviewer coding and data keying errors, changes in interviewers, and imputation procedures. The former type includes respondent bias and variability, which may be affected by length of recall, learning effect of previous interviews, proxy respondents, demographic characteristics, and nonresponse. Each of these possible causes except the last will be discussed briefly here. Nonresponse is discussed in Section III.B.

1. Interviewer Coding/Data Keying

Errors can be made by interviewers and keyers in transcribing the responses in order to produce a computer data file. A monthly verification of SIPP data keying in the regional offices based on a random selection of questionnaires and data fields yields error rates of about .3%. (See, e.g., Linebarger, 1986.) The effect of these errors on reported transitions can only be determined by

examining the individual errors more closely to see if they tend to introduce or mask transitions. If we assume that the interviewer coding rates are of the same magnitude, the overall effect of these sources on the reported patterns is minimal.

2. Change in Interviewer

The respondents in a household become familiar with an interviewer after one or more visits, establishing a rapport that is either beneficial or harmful to accurate response. When a new interviewer arrives the respondents may be more or less willing to reveal receipt of sources such as unemployment compensation. In either case, any change in response would most likely occur for the entire wave, thus introducing false transitions between waves. On the other hand, continuing with the same interviewer may cause under-reporting of transitions.

When new interviewers begin work they do not have the same familiarity with the questionnaire and respondents that more experienced interviewers have. This probably results in some differences in recorded responses, but it is difficult to quantify. The extent of this problem could be investigated by comparing the proportions of between wave transitions reported with the same and different interviewers, as well as with new and experienced interviewers.

3. Imputation

Imputation is used to provide values for items missing from an interview, which usually occurs simultaneously for all four months of a wave. As an example, incorrect imputation of receipt would cause transitions to be recorded when they did not happen, or vice versa. An examination of four waves of data has shown that the proportion of between wave transitions is higher for records with at least one of the waves having imputed data than when both are observed (Weidman, 1987). (See the next section for a more complete description of this work.) However, the nonimputed transitions also exhibit the problem pattern. Thus imputation magnifies an already existing problem.

4. Questionnaire Wording/Design

There are many aspects of the questionnaire and the interview process that affect errors. One general issue is the amount of effort made by respondents and interviewers to provide accurate data. On an interest amount question, for example, at one extreme a respondent might give a top-of-the head guess rounded to the nearest hundred dollars. At the other extreme, a respondent might thoroughly check their records, do some computa-

tions, and add interest across different accounts. How a respondent answers between these extremes is a function of many things, including the specific questions asked, to what extent the questionnaire and training encourage interviewers to probe and to ask for record checking, and the length and complexity of the interviews as a whole.

Another area of concern is the month(s) of receipt for income. Sources of income, assets, etc. received at some time during the wave are determined in the interview before the actual months of receipt are. During the probe for sources, the respondent may forget (or not consider important) a source that was received in only one month of a wave, the interviewer or respondent may lack an understanding of the correct source and misreport it, or the respondent may answer without thinking. These and other sources of response variance are related to the questionnaire format.

The specific months of receipt for each source of income, assets, etc. are determined later in the interview when the amounts are recorded. The months of receipts are queried for beginning with the last month of the wave. If this query began with the first month instead, the respondent might think more carefully about the actual months of receipt and avoid some of the above problems, because a longer recall would be required immediately. This could be a major cause of the observed pattern of transitions, since many people are affected in the same way by the questioning.

5. Length of Recall

This problem is related to the queries about specific months of receipt of sources proceeding from the most recent to the most distant month. A person may report a transition in the wrong month by not remembering the exact month of occurrence. It may be easier to report the receipt state as being the same for all four months in a wave than trying to remember whether it changed 3 or 4 months ago, or if the receipt state in the first month was different than in the other three months the respondent may forget it.

6. Learning Effect

After one or more interviews a respondent may determine that a receipt="yes" requires more additional questions than does a receipt="no". This would lead to excessive between wave transitions from receipt to nonreceipt. At a later time point a person may begin receipt and not report it for this same reason. This would lead to too few transitions from nonreceipt to receipt being reported regardless of the month in which they occurred.

7. Proxy respondents

Changing between proxy and self response may cause reported transitions that did not occur or misplace their month of occurrence. If the change is from self to proxy to self in successive waves, then errors in reporting by the proxy can be corrected through the source roster questions. However, if the proxy response continues this correction will probably not occur. Within wave transitions may be omitted or misplaced because of inadequate knowledge of the proxy.

Weidman (1986) has shown that proxies report a smaller percentage of receipt for many sources than do self respondents. This may cause errors in both between and within wave transition counts. However, there could be legitimate causes of this result other than proxies lacking knowledge about the missing respondents. A further investigation of the characteristics of proxies is required, but because the proportion of self respondents is so high, these errors can only be a minor cause of the observed pattern.

8. Demographics

It may be that respondents with certain combinations of demographic variables report a smaller proportion of receipt of certain sources than actually occur. Identification of such effects would allow us to adjust the data to allow for them or to alter the questionnaire in order to improve respondent accuracy. An investigation of certain demographic variables was made and showed only small effects of some combinations for some sources (Weidman, 1986).

B. Nonresponse and Coverage of Population

Knowledge of rates and causes of nonresponse is important in evaluating the quality of SIPP. This section discusses SIPP nonresponse rates and compares them with those of other surveys. Before discussing this in detail, it is worth mentioning various type of nonresponse.

Every household survey includes individuals who do not respond or respond partially to the questions posed. This nonresponse can be divided into the following categories:

Household Nonresponse: Every member of the household is a noninterview.

Person Nonresponse: A member of an interviewed household could not be interviewed and a proxy interview is not obtained. It is called a type Z noninterview.

Item Nonresponse: A response to a given question is not available.

Table 16 presents response rates for the 1984 SIPP Panel, the National Medical Care and Utilization Expenditures Survey (NMCUES) and the PSID. These rates are not directly comparable due to differences in contents of the surveys, recall periods, frequency of interviews, etc. However, they do provide a general idea about the range of person response rates in multiple interview surveys.

Ongoing statistics have been kept on the distribution of noninterviews and their causes. There are 32,985 persons who were interviewed in wave 1, did not leave the universe, and were not cut from the sample. 69.8% of these were interviewed in each wave through the eighth and 20.2% became and remained noninterviews (including missing both waves 7 and 8). The importance of adjustment becomes important when this attrition is taken into account.

Dahmann and McArthur (1987) studied all persons at least 15 years old who were interviewed in the first wave and survived the fifth-sixth wave sample cut. They looked at differences in characteristics between persons with different interview response patterns. One of the comparisons was between people who responded in all waves and those who were missing at least the last two interviews. Persons who left the universe were not included in these calculations. For each of 23 variables recorded in the first interview, the distributions of these two groups were compared using chi-square tests adjusted by a factor of 3 to take account of the sample design. Significant differences at the 10% level were detected for most of these variables: regional office, size of SMSA, ownership of living quarters, interview status, length of interview, relationship to reference person, household size, age, sex, race, ethnicity, mover status, marital status, hours worked per week, employment status, household and person monthly income, having savings account, and having other types of assets.

McArthur and Short (1986) looked at the relationship between changes in these characteristics at an interview and whether or not a person became a noninterview for the next interview and all interviews through the fifth. There appeared to be relationships for changes in the number of persons in the household, employment status, household income and residence. The results of these studies have led to further work which is currently being pursued. That is, what combinations of variables differentiate persons who become and remain nonrespondents, and what variables and responses at one interview are related to a person becoming a nonrespondent at the following interview? It is hoped that the results of this work will lead to improved adjustments for nonresponse.

Item nonresponse rates for asset amounts were compared for the SIPP and the ISDP in the Current Population Reports, Series P-70, No. 7. It shows that SIPP item nonresponse rates are very large for some items such as value of own business (38%) and market value of stock and mutual fund shares (41%), but they are significantly lower than the ISDP rates for all the items.

Table 17 presents overall item response rates in the SIPP and the CPS for selected income types. These rates for the SIPP are based on core data. The overall item response rate is derived based on household, person and item nonresponse rates. These overall item response rate (100-nonresponse rate in %) for the SIPP are lower than for the CPS for all items presented in the table.

Undercoverage in a survey has an adverse effect on the quality of survey estimates. As a part of the evaluation of the SIPP data quality, the SIPP coverage of the target population by age, race and sex was examined. (Coverage is the ratio of the SIPP estimates of number of people in a specific demographic group to the corresponding independent estimate. Note that the SIPP estimate used is after adjustment is made for noninterviews. This adjustment increases the estimates according to the number of noninterviews, and therefore the indicated undercoverage is not explained by noninterviews. Also, the independent estimates are updated 1980 Census figures, without adjustment for Census undercount. Undercoverage is worse when Census undercount adjustment is included.) The examination showed that, like other household surveys, the SIPP also has a differential coverage by age, race and sex. The coverage ratios for the SIPP and CPS are about the same and are lower for blacks than whites, lower for males than females and are worst for black males 22-24 years of age in both surveys. As examples, SIPP undercoverage as compared to the Census is about 7% for nonblack females and about 15% for Black males.

Nonresponse and undercoverage in surveys are compensated for by complex imputation and/or weighting procedures. These procedures are developed on the assumption that within a demographic group, the persons who respond are similar to those who do not respond. In real life this is not true. Therefore, the quality of the survey estimates including estimates from the SIPP is affected adversely due to lack of complete coverage and nonresponse, and biases exist in estimates to the extent that persons in missed households or missed persons in interviewed households have different characteristics than the interviewed persons.

C. Examination of Error Sources

Several studies at the Census Bureau have examined one or more of the error sources identified in the previous section. In this paper we summarize the results of four of them. They

include a brief look at recall lag, a look at some possible causes of observed transition patterns, an examination of some possible causes of attrition, and an approach to modeling respondent error. The first two of these are presented here, the third in the previous section, and the last in section IV.

The recall effect study (Petroni, 1986) used data from September 1983 to attempt to determine if the number of months between occurrence and reporting of an event affects the reported value. For individuals three benefit sources, labor force activity and monthly income categories were tested. Eight benefit sources and monthly income categories were tested for households. Only one of twenty categories tested significant for recall lag effect at the .05 level, using chi-square tests adjusted for weighted data. This lack of recall lag effect is supported by examination of the data performed as part of the second study. There were extremely few cases where a change in receipt status was reported as occurring within a wave for the several income sources examined. This indicates that for many questions respondents give the same response (perhaps the current state) for all four months of a wave and thus only report changes at the beginning of a wave.

The transition pattern study (Weidman, 1986) examined three possible causes that could contribute to the reported between/within wave pattern of transitions for eight income sources: demographics, interview status (self or proxy respondent), and imputation procedures. We give a brief description of this study and its results.

The income sources examined were social security, unemployment compensation, private pensions, VA compensation and pension, supplemental security income, child support and AFDC. Demographic characteristics that were examined as possible causes of the reported patterns were age, sex, race, marital status, education, relationship to principal person, household size, tenure, and standard metropolitan statistical area (SMSA) size. The distribution of gross flows in receipt status between consecutive months for each income type was computed with respect to all pairs of demographic characteristics and interview status. There are four possible gross flow states for each pair of consecutive months: RR, RN, NR, and NN, where R=receipt and N=nonreceipt. RN and NR denote transitions between receipt states.

In light of the patterns reported by Burkhead and Coder (1985), how is it determined if any relationships exist? For any combination of demographic variables to be a determinant of this change, we would have to observe a huge difference in the number of transitions reported in the first month of a wave as compared to the last three months, but a much smaller difference for other combinations.

Within each cell defined by a particular pair of demographic characteristics, we calculate the probability of each receipt state, P_iAB = P(receipt state AB for cell i). Let P_iAB_w denote such a probability within waves and P_iAB_b the corresponding between wave probability. Compare P_iNR and P_iRN for between waves to those for within wave. If this demographic combination has no relationship to gross changes, the ratios P_iNR_b/P_iNR_w should be fairly constant for all i, as should the ratios P_iRN_b/P_iRN_w. If one and/or both of these sets of ratios differ "greatly" between cells, this indicates the type of relationship we are looking for.

For the second part of this study there are four possible interview statuses of interest for two consecutive months: SS, SP, PS, and PP, where S=self and P=proxy. When examining interview status the situation is somewhat different than for combinations of demographic characteristics. This is because two of the interview status pairs, PS and SP, cannot occur within waves. In this case we look for large differences in the distributions of P_iNR_b and P_iRN_b between cells.

In either case we must be careful about looking at differences for probabilities based on very small numbers of observations because of the resultant large variances in proportions. We present two pairs of tables to represent the results of these comparisons. Tables 18 and 19 give the results for food stamps for sex by interview state. Tables 20 and 21 give the results for food stamps gross flows computed for race by sex. These tables are typical of the results obtained.

A result was noted for interview status, although no major influences on the reported pattern were identified based on the ratio and probability comparisons. For food stamps and social security, larger proportions of receipt of sources were reported by self-respondents than by proxies. Also, there is usually a higher proportion of transitions between waves when at least one of two consecutive months has a proxy response than when both of the months are self-reported.

In the last part of this study the proportion of gross flows that were transitions were calculated for consecutive months without imputation and with imputation. (See tables 22 and 23.) They show a larger proportion of between wave transitions when at least one of two consecutive months is imputed than when both of the months are reported. It may be that people with transitions are more likely to be nonrespondents, so we should not reach any conclusions regarding imputation without a closer examination of the data.

IV. HOW ESTIMATES CAN BE IMPROVED

In this part of the paper we briefly discuss a number of research areas. The first set of 12 topics use general research

to improve our knowledge in some aspect of SIPP quality. The second set of 14 topics goes further in that the research is intended to lead to changes that would improve quality. This is of course not a complete list of possible research, but we have attempted to be fairly comprehensive, possibly including some topics that are not very promising.

Due to limited resources, we anticipate doing work only in a few of these areas at the Census Bureau, and thus strongly encourage others to also work in these areas. We would be happy to talk to anyone with ideas for one or more research projects they would like to conduct.

A. Research for Improved Understanding

1. Time-in-Sample Bias

A very little information about this bias is available from a single study (Coder, 1987a) using only a limited amount of SIPP data. It is generally important to know how large this bias is. In particular, a suggestion has been made to have only one panel in the field at a time. Thus, in one year all addresses would be in their first set of interviews and in the following year would be in their second set of interviews. This is an attractive idea if there is little or no time-in-sample bias but has obvious major problems if bias is high.

2. Improvement of Independent Estimates

For several types of income, SIPP estimates of number of recipients and of amount have been compared to other estimates such as from the Bureau of Economic Analysis (BEA) and the Social Security Administration (SSA). As discussed earlier in the paper, these comparisons generally show SIPP estimates as too low, sometimes by small amounts and sometime by large. The independent estimates are usually for a slightly different universe, use slightly different definitions for the income source, and are subject to some biases of their own. Thus, especially for income sources where SIPP estimates are only a little lower, it is not clear if SIPP is underestimating recipients and amounts. Investigation into the independent sources could be done. For example, we may be able to adjust some BEA estimates for definition differences in some income types. In some cases, such adjustments have already been made to independent source estimates, but they were prepared in 1979 and may be out of date for the purpose of comparison.

3. Recall Errors

The only investigation of recall errors used September 1983 data (Petroni, 1986). That month was in the first wave of the survey and may not be representative of other

waves. Thus, a series of comparisons should be made, including comparisons for population subgroups. Better knowledge about recall errors will be particularly needed if the reference period is lengthened.

4. Direct Analysis of Gross Flow and Spell Data

The simplest form of analysis is subjective analysis of gross flow and spell data. One looks for illogical patterns and anomalies and postulates possible or likely causes for problems found. Much work of this type has of course been done (see, Burkhead and Coder (1985) for example), but more could profitably be done.

A follow-up to this subjective analysis is to identify individual cases where incongruous situations occur and then carefully examine the questionnaires to try to understand what might have happened. Examples of incongruity are no increase in social security income at a time when a cost of living increase in benefits occurs (Kalton and Miller, 1986) or a pattern of frequent changes in receipt/non-receipt for an income source. Little of this type of analysis has been pursued.

Another relatively simple type of analysis is the comparison of gross flows within an interview period to those between interview periods. As discussed above, this has already been done for a number of characteristics, but it could be done for many more characteristics for the 1988 panel to understand effect of changes in the questionnaire.

5. Response Variance Estimation from Reinterview

The reinterviews conducted in SIPP allow for estimates of response variance. Simple estimates of response variance can be made for status characteristics which are used to produce gross flow and spell estimates. One would anticipate some large response variances for characteristics for which the seam flows are much greater than the non-seam flows.

Of greater potential value, however, is a detailed analysis of response variance by demographic characteristics and survey procedures. For example, one can compare response variances for different kin relations (head of household, spouse, and other relative), different ages, and self vs. proxy response on both original interview and reinterview. This type of analysis can indicate that problems exist in only certain situations, e.g., response variances are low for self reporters or for some age groups. O'Muircheartaigh (1986) did exactly this type of reinterview analysis for the Current Population Survey (see especially sections 4 and 5 of his paper). Note, however, that caution must be used in drawing conclusions

because of weaknesses in reinterview data and because there is no experimental control over items like self response versus proxy. Again, see O'Muircheartaigh (1986).

In principle, this analysis could be done with already collected SIPP reinterview data. There are however, three major problems. 1) All reinterviews have been done with reconciliation. It has been well documented (see U.S. Census Bureau (1968, p.25) that the estimated response variance in CPS is much lower with reconciliation than without. The reconciliation estimates are believed to be substantially underestimated. 2) Only a small proportion of all the questions have been included in reinterviews, and thus there is only limited data to analyze. Thus, to get a lot of value from this type of analysis, changes will be required in the reinterview program (see 7. below). 3) Reinterview questions are generally incomplete, i.e., reinterview asks only about receipt during the last four months without asking about specific months.

6. Response Variance Estimation Without Reinterview

a. Use of Single Rotation Groups and Reference Months

A proposal has been made to estimate response variance in SIPP without use of reinterview data. Judkins (1985) suggests a complex estimator based on squared differences for single rotation groups and single reference months. The proof that the estimator is an unbiased estimate of response variance requires the assumptions that length of recall does not affect response bias, that response error is perfectly correlated within wave, and that response error is uncorrelated across waves. Though none of these probably hold exactly, they may be close enough to provide useful response variance estimates.

b. Modeling

Another possible approach is to model the distribution of gross changes using either multivariate normal or logit models (Weidman, 1986). For CPS, it has long been known that there is a relationship between the responses to a question and (i) the amount of time that has elapsed between the month of interest and the month of interview, and (ii) the length of time a person has been in the sample. Work on SIPP has shown a relationship of certain self and proxy responses with interview status. Models were proposed for gross flows that make use of similar relationships.

The dependent variable of interest for a given income type is the receipt state identified with the second of two consecutive months. The possible receipt states for month t are (1)RR, (2)RN, (3)NR, (4)NN. Let $y_{ijkt(m)}$ be the number of responses in receipt state m in month t where

i = number of times a person has been interviewed,

j = number of months between month t and month of interview,

k = interview status for months t-1 and t; PP,PS,SP and SS with S=self, P=proxy.

Then the vector $\underline{y}_{ijkt}$ =

$(y_{ijkt(1)}, y_{ijkt(2)}, y_{ijkt(3)}, y_{ijkt(4)})$,

represents the gross flow counts for the combination ijkt.

(i) Multivariate Normal Models. Since the $\underline{y}_{ijkt}$ are vectors of counts, they have a multinomial rather than a multivariate normal distribution. But because of the large sample sizes on which they are based (the total number of counts in $\underline{y}_{ijkt}$), they have that distribution asymptotically. We propose a multivariate analysis of variance (MANOVA) model of the form:

$$E(y_{ijkt(m)}) = \mu_{(m)} + N_{i(m)} + M_{j(m)} + S_{k(m)} + NM_{ij(m)} + NS_{ik(m)} + MS_{jk(m)} + \hat{a}_t \qquad (1)$$

where the terms are

N_i = interview i,

M_j = months of recall between month of interview and month of occurrence,

S_k = interview status,

NM_{ij}, NS_{ik}, MS_{jk} are interactions of these effects, and

$\hat{a}_t$ = month t.

(ii) Polytomous Logit Models. Alternatively, the probabilities of the receipt states could be estimates using logit models. In this method, the likelihood function is the

product of terms of the form

$$P(\underline{Y}_{ijkt(m')}) = (\underline{X}'_{ijkt}\,\hat{\underline{\beta}}_{m'}) \Big/ \sum_{m=1}^{4} \exp\,(\underline{X}'_{ijkt}\,\hat{\underline{\beta}}_{m}).$$

Here $\underline{X}_{ijkt}$ is a vector of 0-1 variables that indicate which main effects and interactions are present for a particular ijkt combination (as in the right hand side of (1)). Thus, we only need the $\underline{Y}_{ijkt}$ in order to determine the likelihood function and the resulting maximum likelihood estimates $\hat{\underline{\beta}}_{m}$.

When using either of these methods, tests for main effects and interactions being zero would be carried out in order to determine which of them influence the reporting of changes in receipt state. There are some technical difficulties that must be addressed when using either of these models.

7. Expanded Reinterview

It is desirable to keep the respondent burden to a minimum for a complex and lengthy survey like SIPP. Therefore, the reinterview program for the SIPP was designed to discourage fabrication of interviewing and to identify those interviewers who fabricate data. The program is very successful in achieving its goal. Unfortunately, it does not provide a good measure of response variance. Considering the problem with gross flows, it is important to explore all avenues that could help in improving these estimates even if it increases respondent burden and the risk of higher nonresponse in subsequent interviews.

As a starting point, the reinterview program could be expanded to measure response variance for selected items. These items may be selected only from one or two sections of the SIPP questionnaire. When sufficient data are available for these, we could replace them with another set of questions to provide response variance measures for items in another part of the questionnaire. This approach does not attempt to provide the response variance for all estimates at the same time and in a short period. However, it does provide valuable information while still keeping the respondent burden moderate and hence minimizing the risk of increasing nonresponse in subsequent interviews.

Beyond a simple expansion, the reinterview could be used as the vehicle for various experiments.

8. Use of Administrative Records

Administrative records could be very useful in increasing understanding in order to improve estimates of gross flows and length of spells. The administrative records could be used at the macro or micro level.

At the macro level, studies similar to validation of food stamp turnover (Judkins, 1986), AFDC turnover (Maher, 1987b) and supplemental security income (Maher, 1987c) would provide information on the quality of additional transition estimates at the macro level. Transition and spell estimates for longer time periods should also be evaluated to assess their quality.

To make the best use of the SIPP, it is extremely important to utilize micro level data. The gross flow estimates suggest problems with the data at the micro level. A micro level match of SIPP data with administrative records has begun at the Census Bureau (Singh, 1986 and Moore, 1986). This study plans to evaluate the SIPP data by matching individual records on recipiency of nine government transfer programs in four states - Florida, New York, Pennsylvania, and Wisconsin and develop a model of SIPP response and imputation errors in measures of program participation and amount received (Moore, 1986). This is a good step in the right direction, but more efforts are needed to evaluate and develop models for other characteristics and/or other states.

9. Special Samples With Known Income Sources

The preceding section discussed getting information on reporting errors through matching of survey data with administrative data. One can also select particularly interesting cases from administrative records to include for evaluation purposes in the SIPP. We might, for example, select some households with multiple recipiency of income/program sources that occur infrequently, e.g., supplemental security income and unemployment compensation, to explore whether we particularly tend to get reporting errors in such cases. We could also plan special reinterviews for households selected from administrative data when sample and administrative records data disagree. No plans for this type of research have been made.

10. Cognitive Research

Cognitive research can be important in a number of areas. Research would be intended to examine cognitive processes of respondents during interviews, to explore outside influences affecting respondent behavior, and to develop improved questions, procedures, etc. Areas of application include coverage problems (especially for Black and

Hispanic males), timing of events (gross flows) and respondent willingness to participate and to consult records.

One way to obtain information is through debriefing of respondents. A debriefing of some respondents after completing all SIPP interviews was done in a reinterview in 1987 (Matchett, 1987). Respondents were asked why they continued to participate and whether they had comments to improve data collection. Analysis is continuing, but some preliminary information is already available. The main reasons for participation are wanting to be sociable, liking the interviewer, and having nothing to hide. Further debriefing should be done, correcting some problems discussed in the initial debriefing, using open-ended instead of fixed response questions and addressing different problems.

11. Basic Coverage Research

As previously discussed, the SIPP and other demographic surveys have much worse coverage than the Decennial Census. One partial explanation is that the Census includes a number of erroneous inclusions, such as duplicates, that are not included in the SIPP. The project here would be to adjust the controls used in forming coverage ratios by excluding the erroneous inclusions. Analysis of such ratios by age-sex-race would improve our knowledge about differences between the SIPP and Census coverage.

Another area of research involves comparisons of survey and Census tabulations. Valentine and Valentine (1971) concluded from a small-scale study on one area that most of the omitted Black males in Census Bureau surveys are household heads. Since the Census has much better coverage than our surveys, the Valentine hypothesis would lead us to expect some significant household composition differences between the Census and our surveys. To examine this, we would compare April 1980 Current Population Survey (CPS) tabulations to Census tabulations. We would use special CPS tabulations that exclude the normal ratio estimation to population control figures.

12. Imputation

Not much is known about the accuracy of SIPP imputation. The imputation may be overcompensating or undercompensating for nonrandom differences, if any, between respondents and non-respondents. Also, the frequency of transitions for imputed cases is much greater than for non-imputed cases for many income sources, suggesting possible deficiencies in the imputation methods (see tables 22 and 23). Also, persons who are nonrespondents because they move to an unknown address appear to have different

characteristics than other non-respondents. Thus, it may be that adding a variable about movers would improve the imputation system. In general, research is needed into how well the imputation system is working.

B. Research for Improving Estimates

1. Reducing Complexity

There are 3 panels in SIPP from February through August and 2 panels from September through January. This makes for a variable workload, resulting in some regional office clerks working only part of the year on SIPP and in difficulties for interviewers. More importantly, each panel has a somewhat different questionnaire, so that interviewers have to deal with up to 3 different questionnaires at a time. This necessitates multiple clerical and supervisory procedures. Training is made more difficult.

If the SIPP questionnaires were short and simple, having 3 versions would be less of a problem. But the basic questionnaire is complex and requires considerable interviewer knowledge in order to administer it correctly. As an example, interviewers must know the difference and distinguish in the interview between a bank certificate of deposit and a statement savings account to collect data of good quality.

It is believed that questionnaire length and complexity, together with having as many as 3 questionnaires simultaneously in use, results in interviewing errors, less probing than desired, and infrequent checking of records for income amounts.

There are several things that would reduce complexity. First, we could redesign SIPP so that only 1 or 2 panels would be interviewed at a time. Four such options have been mentioned. The simplest of these options would have each panel in sample for exactly 3 years and a new panel would be introduced only once every 3 years. Its main disadvantage is that comparisons of estimates would be adversely affected by time-in-sample bias. The other 3 options have new panels introduced at one to two year intervals. They would be less affected by time-in-sample bias, but would have 2 panels being interviewed simultaneously all of or part of the time.

A second way to reduce complexity is to shorten the core questionnaire. A major decrease in length could help substantially. Interviewers would have less to learn and remember, and shorter interviews would be conducive to more probing, more use of records by respondents, and higher response rates. Of course, a major disadvantage is less data and information from the survey.

Another related way to reduce complexity is to reduce the number or/and size of topical modules. This would have the same advantages and disadvantages as would shortening the core questionnaire.

2. Improving Field Procedures

Beyond initial training, interviewers are monitored through observation, reinterview, and administrative data. Periodically all interviewers are observed by their supervisor or a Supervisory Field Representative. The interviewer receives positive and negative feedback, as appropriate during the observation, and further action is taken if serious problems are uncovered. Reinterview is used primarily to ensure that interviewers do not fabricate interviews. Interviewers are informed about the reinterview results. Finally, data are kept on productivity and noninterview rates. Appropriate action is taken when there are indications of low productivity or high noninterview rates.

Over the last year or two, significant improvements have been made in the monitoring programs. Through the use of microcomputers and data base systems, historical data on interviews is much more readily accessible to the supervisors. There have been changes towards more positive feedback to interviewers. Previously, somewhat rigid standards for acceptable interviewer performance have been changed to flexible guidelines, with emphasis on supervisors making their own decisions on when an interviewer has a serious performance problem that requires corrective action. However, further improvements are still needed. Supervisors need more training on how to use the data available to them for evaluation and coaching. There is still a need for more communications, especially positive feedback, by supervisors.

3. Improving Training

Training is particularly important in SIPP since it is such a complex survey. Holt (1986) has made some specific recommendations for improvements in training that should be pursued. The Bureau is currently evaluating these recommendations for possible implementation.

4. Reducing Nonresponse

A gift experiment was conducted on the SIPP 87 panel to see if it reduces nonresponse in SIPP. According to the experiment, a token gift of solar calculators was given to those households who were eligible for interview in April 1987. The complete results of this experiment will not be available until after the panel retires. Three additional ideas are presented below.

First, there can be follow up experimentation to the earlier discussed experiment in which calculators are given to respondents. This would involve different gifts or multiple gifts, or gifts given at different times in the interview cycle.

Secondly, thank you notes handed to respondents at the end of an interview might improve cooperation in future interviews.

Third, providing interviewers and respondents with more information on the survey objectives may be helpful, although some of this is already being done. This would address interviewer observations that some respondents have stopped participation because they don't see a need to answer the same questions over and over again.

5. Dependent Interviewing

Asset and liability questions are asked in the seventh interviews. During a feedback experiment in 1986, some seventh interview respondents were given information on their wave 4 responses. Analysis is still continuing, but preliminary results do not show any evidence of feedback affecting the data (Lamas and McNeil, 1987). Nonetheless, feedback and/or more dependent interviewing may still have potential. For example, Coder (1987c) has suggested that when there is an indicated transition from recipiency to nonrecipiency at the seam, the respondent could be asked how many months it was since the last receipt of that income source. If the answer is not 4 months, the transition may not really have occurred at the seam. Even a different type of feedback on assets and liability might show improvements. Thus, additional experimentation with dependent interviewing would be worthwhile.

6. Reference Period

Various studies (for example, Kobilarcik, et. al. 1983) have shown that the length of recall affects the data quality. As the length of recall varies, the quality of data varies. A better understanding of the gross flow estimates will help in identifying important estimates with large problems. For these estimates, a shorter reference period would be desirable. On the other hand, a longer reference period could be used for items with small problems. However, consideration to the importance of these items needs to be given in deciding the length of the reference period. One suggestion is to have various (differential) recall lengths for different core questions during the same interview. The topical modules already have differential (mixed) reference periods. The mixed reference period approach has also been used for the Consumer Expenditures Survey.

Another suggestion involves frequent brief telephone interviews interspersed with less frequent full interviews. For example the basic interviewing frequency could be increased from 4 to 6 months (with a reference period also of 6 months). In addition, there could be one or two short telephone interviews between the full interviews. The telephone interviews might only ask whether there have been any changes in recipiency status or amounts for types of income.

The main potential advantage is that the 2 month recall would result in more accurate transition data and greatly reduce the seam effect. On the other hand, it is unknown whether such a methodology is feasible; there are several potential disadvantages, and details of the methodology have not been determined.

7. Reducing Response Variance

For transitions that have particularly high response variances, specific efforts can be made to reduce the response variance. In particular, attempts can be made to determine improvements in the questionnaire and/or in the data collection procedures. Proposed methods can then be compared with present methods in experiments that use carefully conducted reinterviews to measure the response variances. This type of undertaking has been started for the American Housing Survey (Schwanz, 1986).

8. Improving Transition and Spell Estimates

There is interest in pursuing any procedural, design or questionnaire changes that could lead to improved transition and spell estimates. One such change that could possibly improve estimates of transition from nonrecipiency to recipiency is to reverse the order in which months of recipiency are asked. Recipiency in the most distant month would be asked first and the most recent month last.

Another potentially helpful procedural change is to present respondents with calendars or diaries that they can keep and use to record relevant dates and income amounts.

Changes can be made for programs that have cost-of-living increases at fixed times during the year. For example, food stamp increases occur in July and October. Reports of such increases could be improved by reminding recipients of cost-of-living increases in the appropriate months.

One suspected cause of false transitions at the seams is inconsistent classification of income sources between interviews. For example, in one interview a respondent

may report Aid to Families With Dependent Children (AFDC) income and the next interview General Assistance (GA) income, whereas in reality the income source was unchanged. The inconsistencies could be reduced if recipients were reminded of some characteristics that uniquely identify a particular program (such as color of check, date mailed, or where it is mailed from). Also, a program edit that was developed for the Income and Survey Development Program (ISDP) to reduce misclassification between AFDC and GA income could possibly be used in SIPP. The ISDP edit "corrected" classifications based on respondent reports on monthly payment amount, unit size, state of residence, WIN participation and Medicaid coverage. The weakness to this edit is that actual survey answers are changed, some of which may have been correct.

9. Increasing Respondent Effort

Improvement of respondent effort could improve data. We could stress to respondents that it's important to us to know the exact months of recipiency, and could ask respondents to make a commitment to answering the questions as well as possible and to think about their answers.

10. CPS Gross Flow Conference Proposals

At the Conference on Gross Flows in Labor Force Statistics organized by the Census Bureau and the Bureau of Labor Statistics, several methods for adjusting for errors in transition estimates were presented. Two papers, Fuller and Chua (1985) and Poterba and Summers (1985) present reasonable and viable adjustment procedures for response error, using reinterview data for estimating response errors. (See also Fuller and Chua, 1986.) Abowd and Zellner (1985) also present a viable procedure which adjusts for missing data (nonresponse in one interview or non-match between interviews) as well as response error. Any of these three procedures could be applied directly to SIPP transition estimates with the availability of estimates of response variance from reinterview or other sources.

The main research required at this point is an in-depth comparison of the three methods, both theoretical and empirical, which might result in one or more new procedures which combine their best features. The goal of such research would be to determine the 'best' adjustment procedure for SIPP transitions. One problem that at least some of the present adjustment procedures have and that needs to be addressed is that adjustment yields negative transitions in some situations. In practice, research is likely to conclude that at least two different adjustment procedures are about equally good. If two or three "best" procedures result in substantially dif-

fering transition estimates from each other, it will be impossible to have much confidence in adjusted transition estimates even if there is a consensus that the adjusted estimates are better than the unadjusted.

11. Imputing versus Weighting Adjustments

How to handle missing data for longitudinal analysis is an important issue, especially when the sample unit is a noninterview for only some of the interviews. Kalton (1986) discusses various alternatives to deal with such situations. The preliminary evaluation of the missing wave data for wave 8 suggests that, in certain situations, imputation could be used with little affect on gross flow estimates (Huggins, 1987). However, more research should be performed in deciding when and which of the two procedures should be used.

12. Improving Wage and Salary Income

One possible problem contributing to wage and salary income underestimates is that some respondents report take-home pay instead of gross pay (Coder et. al. 1987). One possible improvement may be to ask for both take-home pay and gross pay.

13. Improving Interest Income

SIPP clearly underestimates interest income recipients and amounts. There are several ways to improve the reporting of interest income.

One approach is to use IRS records instead of respondent answers, although this may make subannual estimates impossible. Since interest income data on IRS records is not available by source, this approach has the potential to improve only an estimate of aggregate interest income for federal tax filers. Another approach is to give respondents a notebook in which to record the information. Perhaps the notebook could be made useful for other things as well, and so function as a token reward for cooperation. A third approach is to provide more training to interviewers on the various sources of interest income so that interviewers might more effectively probe. A fourth approach is for respondents to tell us the principal and interest rate for each source of income rather than the amount of interest.

14. Improving Child Care Questions

In the child care topical module, questions are asked about child care arrangements. Among other things, estimates are produced on the number of children, both young and old, who care for themselves after school while their parents work. We have asked about child care arrange-

ments directly. These questions can be very sensitive for parents whose child care arrangements are not very good for young children, and thus such parents may frequently mis-report on our questions. Research on this may lead to better questions and better data.

15. Improving Assets Data

Obtaining accurate information on assets and liabilities is very difficult for all surveys. Assets is an area where many respondents are leery of providing information or are not knowledgeable. It is possible to get at least some assets data from administrative sources by matching on social security number. However, there are major problems of administrative data not being consistent with survey definitions and categories. The work required to be able to use each data source will be substantial. Thus, we may be able to improve assets estimates by substituting administrative data for survey data.

V. SUMMARY

In this paper we have taken a brief but wide-ranging look at studies that have been carried out to evaluate many aspects of SIPP data quality, and we have proposed additional areas of study aimed at improving and further evaluating data quality. It is not possible to make a general statement about the results of the studies, but we can summarize them for different types of data.

Estimates were classified as belonging to two groups --cross-sectional and gross flow/spell. SIPP cross-sectional estimates of the number of recipients for and amounts received from several government programs by quarter are lower than for administrative sources, but for amounts SIPP's generally higher than for the CPS. However, the number of people receiving and the amounts received for unemployment compensation show a decreasing trend compared to independent sources. Estimates of annual income of various types using the SIPP longitudinal file were comparable for the SIPP and the CPS, but poverty rates are lower for the SIPP and thought to be somewhat closer to the actual because of SIPP's better coverage of transfer program income and shorter recall period.

Estimates of rates of change in table 7 show differences between the SIPP and administrative sources, but only one of them is statistically significant. Comparisons of differences in estimates one year apart of the number of households having certain income sources are statistically significant for 4 out of 5 sources. Further investigation of these differences is needed.

Much work has been done on gross flow estimates because of the observed problem of a large percentage of transitions being reported as occurring between waves. Validation of exit and

start-up rates for food stamps, AFDC, and SSI has produced mixed results for macro level use of the data, suggesting that each benefit source should be individually evaluated. A study of the relationship of demographics, imputation procedures and interview status with this pattern of reporting showed no large-scale results. However some small-scale results indicated that proxy respondents and imputation contribute to overestimates of numbers of transitions between waves. To understand the effect of gross flow patterns on the micro-level analysis, Young (1989) computed correlations between a number of different events and amount change status. Except for correlation of 'marital status' and 'married spouse present' with other characteristics they did not show patterns of distortion in bivariate relationships. However, until more analysis is completed, one should be careful in judging the utility of the data for multivariate analysis at the micro-level.

Nonresponse takes various forms including household, person and item. One serious problem with the SIPP is the number of people who become and remain nonrespondents, approximately 20% of the sample by the eighth interview. A study comparing those who missed the last two waves with those responding in all waves shows many variables related to this nonresponse. Further investigation of this data is being carried out. Item response rates for selected income types are given in table 17 and show lower rates for the SIPP than the CPS.

As this summary indicates, the SIPP data quality compares favorably with other sources in some cases and not so favorably in others. This is not surprising since the SIPP uses such an extensive questionnaire, as well as topical modules, that attempts to collect accurate information for many constituencies. Further studies should be carried out to evaluate variables and error sources that have not yet been treated. In addition, research should be carried out on methods for directly improving the quality of data through better interviewing procedures.

ACKNOWLEDGEMENTS

The authors wish to express their appreciation for the work of Kimberly Wilburn and Cora Wisniewski, without whose dedication and willing attitude this paper could not have been completed. Due to their exceptional efforts and skill; we were able to dispatch a lengthy study to meet an imposed deadline. We would also like to thank Mary Ellen Beach, Jack McNeil, Daniel Kasprzyk, Norman Johnson, and Randall Parmer, all of the Census Bureau, for their valuable comments.

Table 1. Comparisons of Estimated Numbers of Income Recipients and Estimated Aggregate Income Amounts Received for Selected Income Types: SIPP vs Independently Derived Estimates vs the Current Population Survey

	SIPP as a Percent of the Independent Estimates of Monthly Average Recipients for Selected Income Types by Quarter	SIPP as a Percent of the Independent Estimates of Aggregate Income Amounts Received for Selected Income Types by Quarter	CPS (1983) as a Percent of the Independent Estimate Aggregate Income Amounts Received
Wage and Salary			
3rd Quarter 1983	(x)	95.0	99.0
4th Quarter 1983	---	94.3	
1st Quarter 1984	---	93.2	
2nd Quarter 1984	---	94.4	
3rd Quarter 1984	---	95.2	
4th Quarter 1984	---	94.5	
Federal Supplemental Security			
3rd Quarter 1983	92.0	89.8	84.9
4th Quarter 1983	91.3	93.5	
1st Quarter 1984	94.8	96.4	
2nd Quarter 1984	98.2	97.4	
3rd Quarter 1984	98.3	98.6	
4th Quarter 1984	98.1	99.2	
Social Security			
3rd Quarter 1983	99.2	99.6	91.7
4th Quarter 1983	96.3	100.6	
1st Quarter 1984	97.3	100.5	
2nd Quarter 1984	97.7	101.1	
3rd Quarter 1984	97.5	101.3	
4th Quarter 1984	97.5	101.6	
Aid to Families with Dependent Children 1/			
3rd Quarter 1983	78.5	76.2	76.0
4th Quarter 1983	79.2	78.5	
1st Quarter 1984	84.5	85.3	
2nd Quarter 1984	86.0	86.0	
3rd Quarter 1984	82.0	80.2	
4th Quarter 1984	80.7	78.8	
Food Stamps			
3rd Quarter 1983	89.5	90.1	71.2
4th Quarter 1983	91.1	83.1	
1st Quarter 1984	90.8	85.2	
2nd Quarter 1984	90.5	86.2	
3rd Quarter 1984	90.3	84.6	
4th Quarter 1984	91.7	83.6	
Veterans' Compensation or Pension			
3rd Quarter 1983	89.2	78.9	63.3
4th Quarter 1983	89.7	79.9	
1st Quarter 1984	90.6	78.0	
2nd Quarter 1984	90.8	74.5	
3rd Quarter 1984	89.8	76.3	
4th Quarter 1984	93.3	79.7	

1/ The amount excludes dependents covered by payments.

(x) Not Applicable

Table 2. Comparisons of SIPP State Unemployment Compensation with Estimates Derived from Independent Sources

(Monthly Averages for Specified Quarter. Recipients in thousands, aggregates in millions)

Period	SIPP as a Percent of Independent Estimate[1]		CPS 1983 Estimates as a Percent of the Independent Estimates Aggregate Income Amounts Received
	Recipients	Aggregate Amount Dollars	
1983			
Third Quarter	100.9	102.2	75.5
Fourth Quarter	103.4	106.8	
1984			
First Quarter	82.6	85.2	
Second Quarter	82.5	83.1	
Third Quarter	78.5	80.3	
Fourth Quarter	95.1	100.9	
1985			
First Quarter	85.5	94.8	
Second Quarter	77.3	77.7	
Third Quarter	72.8	72.6	
Fourth Quarter	79.1	77.4	

[1]Independent estimates exclude Federal Supplemental Compensation

Source: Coder, J. (1987b)

Table 3. Comparison of Annual Aggregate Income Estimates from the March CPS and SIPP 1983-1984 Longitudinal Research File

(In millions of dollars)

Income source	SIPP 1983-1984	March CPS	
		1984	1983
Cash transfers, total	216,326	200,620	197,975
Social Security	153,958	147,503	138,293
Railroad Retirement	5,603	3,973	3,975
Federal SSI	8,859	8,444	7,647
Public assistance, total	14,643	13,407	12,878
AFDC	11,881	10,972	10,523
Other	2,762	2,435	2,355
Unemployment Compensation, total	14,911	12,169	19,720
State Unemployment Compensation	14,060	(NA)	(NA)
Other	851	(NA)	(NA)
Veterans' Payments	10,978	8,349	8,831
Worker's Compensation, total	7,374	6,775	6,631
"State" worker's compensation	6,041	(NA)	(NA)
Other Compensation	1,333	(NA)	(NA)
Pensions, total	92,619	85,448	79,718
Private pensions, total	40,319	37,266	34,636
Company or union pensions	32,874	(NA)	(NA)
Other private pensions	7,445	(NA)	(NA)
Federal pensions	19,593	17,154	17,720
Military pensions	15,556	15,328	14,095
State and local pensions, total	17,151	15,700	13,267
State	12,201	(NA)	(NA)
Local	4,950	(NA)	(NA)
Interest income	115,687	138,661	118,800
Dividends	38,251	30,657	27,286
Rents and royalties	16,834	17,725	16,483
Estates and trusts	5,085	7,835	6,666
All other income, total	36,720	30,487	27,258
State SSI	101	(NA)	(NA)
Foster child care	207	(NA)	(NA)
Child support and alimony	8,551	9,401	8,323
Income from charity	58	(NA)	(NA)
Money from friends or relatives	6,441	4,757	5,358
Income from roomers or boarders	165	(NA)	(NA)
Financial investments	16,389	(NA)	(NA)
Other income not included elsewhere	4,808	16,329	13,577
Food Stamps	9,267	7,555	7,471

NA Not available.

Source: Coder (1986b)

Table 4. Comparison of SIPP and March CPS Estimates of Persons Ever Receiving Benefits from Selected Programs

Selected income sources	SIPP 1983-1984	CPS	
		1984	1983
Social Security	34,122	32,182	31,731
Federal SSI	3,941	3,568	3,442
State Unemployment Compensation[1]	9,082	7,693	10,109
Veterans' Payments[2]	3,790	2,865	3,156
AFDC	3,987	3,561	3,468
Worker's Compensation	2,329	2,478	2,382
Private pensions	8,499	7,951	7,618
Federal pensions	1,937	1,555	1,609
Military pensions	1,297	1,493	1,337
Interest income	123,135	99,045	99,005
Dividends	26,807	19,858	18,690
Rents and royalties[3]	14,040	12,461	11,836
Estates and trusts	521	1,384	1,239

[1] CPS estimates may include a small number of persons receiving other types of "unemployment" benefits but no State unemployment compensation.

[2] CPS estimates include G.I./VEAP beneficiaries who do not receive cash veterans payments. The SIPP figure excludes this group.

[3] The SIPP estimates excludes persons receiving royalties but not rental income.

Source: Coder (1986b)

Table 5. Comparison of Cross-Sectionally Derived Quarterly Estimates with Fourth Quarter 1983 Estimates Derived from the Longitudinal Research File

(Recipients in thousands. Monthly averages)

Selected income sources	SIPP Cross-sectional estimates				SIPP Q4-83 estimates based on longitudinal file
	Q4-83	QI-84	Q2-84	Q3-84	
Recipients					
Social Security	31,854	32,370	32,432	32,376	31,924
Federal SSI	3,216	3,362	3,492	3,549	3,346
State Unemployment Compensation	2,878	2,982	2,212	1,927	3,013
Veterans' Payments	3,568	3,546	3,503	3,435	3,527
AFDC	2,894	3,129	3,171	2,973	3,065
Food Stamps	6,746	6,917	6,775	6,416	6,916
Median Amount					
Social Security	$385	$398	$402	$402	$385
Federal SSI	209	211	208	206	214
State Unemployment Compensation	400	396	379	361	364
Veteran's Payments	131	126	124	125	120
AFDC	285	289	293	287	285
Food Stamps	99	101	99	96	99
Mean Amount					
Social Security	$395	$405	$409	$411	$396
Federal SSI	216	221	218	218	217
State Unemployment Compensation	414	405	406	395	410
Veterans' Payments	235	229	226	232	236
AFDC	314	316	318	319	319
Food Stamps	111	113	113	111	111

Source: Coder (1986b)

Table 6. Comparison of Mean Annual Income Amounts from the March CPS and SIPP 1983-1984 Longitudinal Research File

Income Source	SIPP 1983-1984	March CPS	
		1984	1983
Social Security	$ 4,512	$ 4,583	$ 4,358
Railroad Retirement	6,448	6,190	6,098
Federal SSI	2,248	2,366	2,221
AFDC	2,980	3,072	3,034
Federal Pensions	10,115	11,032	11,013
Military Pensions	11,586	10,267	10,538
Dividends	1,427	1,543	1,459
Estates and Trusts	9,709	5,660	5,379
Food Stamps	954	1,070	1,042

Note: This limited list of income types includes only those for which directly comparable mean income could be derived given the data available at the time of preparation.

Source: Coder(1986b)

Table 7. Rates of Change in the Number of Program Participants from SIPP and Independent Sources

Comparison*	Characteristic	Estimates from SIPP (84-83)/83	Estimates from Other Source* (84-83)/83	Difference
1	Social Security	.029	.010(A)	.019
2	SSI	.096	.028(A)	.068
3	AFDC	-.018	-.013(A)	-.005
4	Food Stamps	-.050	-.047(A)	-.003
5	Average household income	-.066	.081(C)	-.015
6	Average monthly earnings of married, spouse present, male working fulltime	-.036	.033(C)	-.069**

*"A" stands for the administrative record and "C" stands for CPS.

**Stands for significant difference.

Source: Kim, J. (1985)

Table 8. Differences of SIPP Estimates Between 1983 and 1984 3rd Quarters

Characteristics	NUMBER OF HOUSEHOLDS (1) 1983	(2) 1984	(3) Difference (2) - (1)
Interest generating Assets	57,170,000 (30,278)**	56,249,000 (372,632)	-921,000*
Cash Dividends	11,317,000 (258,430)	10,506,000 (250,301)	-811,000*
Rental Income	6,457,000 (201,253)	5,981,000 (194,254)	-476,000*
Income from Mortgage	2,185,000 (120,079)	2,213,000 (194,254)	28,000
Other Type of Financial Asset	2,319,000 (123,610)	2,019,000 (115,538)	-300,000*

** The number in the parentheses is the standard deviation of the number just above it.

* Indicates that the calculated test statistic is significant at the 5-percent significance level.

Source: Kim, J. (1985)

Table 9. SIPP Asset and Liability Estimates Compared to Federal Reserve Board Balance Sheet Data for the Household Sector: 1984

(Number in billions except for median networth)

Category	FRB balance sheet	SIPP	Ratio of SIPP to FRB balance sheet
A. Equity in owner-occupied housing	$ 2,174.2	$2,823.6	1.30
Gross value	3,482.7	3,958.2	1.14
Debt	1,308.5	1,134.6	0.87
B. Equity in motor vehicles	287.0	410.5	1.43
Gross value	459.6	558.8	1.22
Debt	172.6	148.3	0.86
C. Equity in noncorporate business	2,229.7	1,680.2	0.75
D. Financial assets	3,812.0	2,826.1	0.74
1. Interest-earning assets[1]	3,195.2	1,635.7	0.51
2. Corporate equities[2]	1,456.7	1,062.7	0.73
3. Other financial assets[3]	160.4	127.8	0.80
4. Less: Financial assets held by nonprofit sector or in personal trusts	(840.0)	X	X
E. Installment and other consumer debt[4]	379.9	241.5	0.64
F. Net Worth (A+B+C+D-E)	8,122.9	7,498.8	0.92
G. Median Networth	30,550.0	32,670.0	1.07

NA Separate estimates not available.

X Not Applicable.

[1] Includes passbook savings accounts, money market deposit accounts, certificates of deposit, checking accounts, money market funds, U.S. Government securities, municipal or corporate bonds, saving bonds, IRA and KEOGH accounts, and other interest-earning assets.

[2] Includes equities in stocks, mutual fund shares, and incorporated self-employed businesses or professions.

[3] Includes mortgages held by sellers and other financial assets not otherwise specified.

[4] Excludes debt for automobile and mobile homes.

Source: U.S. Bureau of the Census, Current Population Reports, Series P-70, No. 7, 1986

Table 10. Median Wage and Salary Income in 1984 From the WAVE 6 Topical Module

(Based on unweighted observations)

Record usage and respondent type	Total	Men	Women
Used W-2 Form			
Total	15,222	20,990	10,825
Self	14,422	17,967	11,255
Proxy	17,897	21,031	7,107
Did not use W-2 Form			
Total	11,515	15,963	8,542
Self	11,168	16,009	8,897
Proxy	12,273	15,896	6,632

Source: Coder, J. (1987d)

Table 11. Time-in-Sample by Rotation Covering a Reference Month for SIPP 1984 Panel

Reference Month	Rotation Group 1	2	3	4
November "83	2	2	1	1
December "83	2	2	2	1
January "84	2	2	2	2
February "84	3	2	2	2
March "84	3	3	2	2
April "84	3	3	3	2
May "84	3	3	3	3
June "84	4	3	3	3

Note: The numbers in the table indicate the Time-in-Sample. For example 2 means the second time interviewed.

Table 12. Month-to-Month Changes in Recipiency and Amounts of Food Stamps for Fully-Interviewed Persons Age 15 Years and Older

TYPE OF CHANGE	MONTH-TO MONTH CHANGES										
	1ST TO 2ND	2ND TO 3RD	3RD TO 4TH	4TH TO 5TH	5TH TO 6TH	6TH TO 7TH	7TH TO 8TH	8TH TO 9TH	9TH TO 10TH	10TH TO 11TH	11TH TO 12TH
TOTAL WITH INCOME IN AT LEAST ONE MONTH..	1927	1927	1927	1927	1927	1927	1927	1927	1927	1927	1927
RECEIVED INCOME IN BOTH MONTHS.....	1287	1306	1325	1211	1334	1341	1351	1224	1327	1326	1329
AMOUNT DECREASED BY 75.0 TO 99.0 PERCENT..	4	3	7	12	0	3	6	6	2	4	4
AMOUNT DECREASED BY 50.0 TO 74.9 PERCENT..	6	7	10	31	5	9	10	36	1	7	12
AMOUNT DECREASED BY 25.0 TO 49.9 PERCENT..	12	24	17	56	10	21	22	68	12	13	10
AMOUNT DECREASED BY 10.0 TO 24.9 PERCENT..	12	22	22	89	15	22	24	123	16	20	21
AMOUNT DECREASED BY 5.0 TO 9.9 PERCENT....	12	9	14	58	5	7	13	69	8	7	16
AMOUNT DECREASED BY LESS THAN 5.0 PERCENT.	9	9	9	103	9	13	18	90	2	6	6
AMOUNT DID NOT CHANGE.....	1166	1131	1108	444	1241	1194	1176	505	1262	1207	1190
AMOUNT INCREASED BY LESS THAN 5.0 PERCENT.	0	34	38	149	7	18	16	84	4	11	7
AMOUNT INCREASED BY 5.0 TO 9.0 PERCENT....	6	16	23	64	7	11	8	34	2	7	10
AMOUNT INCREASED BY 10.0 TO 24.9 PERCENT..	17	11	28	76	9	7	20	63	5	19	14
AMOUNT INCREASED BY 25.0 TO 49.9 PERCENT..	6	12	16	49	11	16	16	46	6	10	13
AMOUNT INCREASED BY 50.0 TO 74.9 PERCENT..	11	9	6	31	5	8	8	35	1	6	9
AMOUNT INCREASED BY 75.0 TO 99.9 PERCENT..	5	6	5	10	3	2	5	20	8	3	1
AMOUNT INCREASED BY 100.0 PERCENT OR MORE.	13	13	22	47	7	10	9	45	6	6	16
FROM POSITIVE AMOUNT TO LOSS.....	0	0	0	0	0	0	0	0	0	0	0
FROM LOSS TO POSITIVE AMOUNT.....	0	0	0	0	0	0	0	0	0	0	0
LOSS BOTH MONTHS.....	0	0	0	0	0	0	0	0	0	0	0
FROM RECEIVING TO NOT RECEIVING INCOME.....	44	48	43	177	25	42	45	180	36	39	33
FROM NOT RECEIVING TO RECEIVING INCOME.....	67	62	63	148	49	55	53	139	38	36	45
DID NOT RECEIVE INCOME BOTH MONTHS.....	529	511	496	391	519	489	478	384	526	526	520

Source: Coder (1986a)

Table 13. Start-up and Exit Rates (Percentages) for Food Stamp Participation

SIPP 84 Panel-Reference Month i to i+1 Across All Four Rotations

	1 to 2	2 to 3	3 to 4	4 to 5	Avg.
Start-up Rate	4.9	4.7	4.5	10.9	6.2
Standard Error[1]	.8	.8	.7	1.1	0.5
Exit Rate	3.3	3.5	3.1	12.8	5.7
Standard Error[2]	.7	.7	.6	1.2	.5

Urban Institute data-Calendar Month i to i+1 in 1983

	6 to 7	7 to 8	8 to 9	9 to 10	10 to 11	11 to 12	Avg
Start-up Rate	6.7	6.9	6.1	6.2	6.7	5.0	6.3
Standard Error[1]	.6	.6	.5	.5	.6	.5	.3
Exit Rate	7.3	5.8	6.7	7.0	6.1	5.1	6.3
Standard Error[2]	.6	.5	.6	.6	.5	.5	.3

[1] For individual pairs of months, a design effect of 1.8 is assumed. For the average, a design effect of 2.6 is assumed to reflect the correlation between the individual pairs reduced by being in the same set of PSUs. The monthly sample sizes were around 1350. For the average, the sample size is quadrupled.

[2] For individual pairs of months, a design effect of 1.3 is assumed. For the average, a design effect of 2.0 is assumed. The monthly sample sizes were around 2600. For the average, the sample size is to be sextupled.

Source: Judkins (1986)

Table. 14 Start-up and Exit Rates (Percentages) for AFDC Participation

SIPP 84 Panel-Reference Month i to i+1 Across All Four Rotations

	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	1-5 Avg.	5-9 Avg.	1-9 Avg.
Start-Up Rate	2.9	2.7	2.3	9.6	2.9	2.1	1.6	10.2	4.4	4.2	4.3
Standard Error[1]	.9	.9	.8	1.6	.9	.8	.7	1.6	1.3	1.3	1.5
Exit Rate	1.4	1.9	1.5	8.1	1.0	1.4	2.1	9.9	3.2	3.6	3.4
Standard Error[1]	.6	.7	.7	1.5	.5	.6	.8	1.6	1.2	1.2	1.4

AFDC Quarterly Averages

	Quarter 3 1983	Quarter 4 1983	Quarter 1 1984	Quarter 2 1984	July-Dec.83 Avg.	Oct.-June(83-84) Avg.	July-June(83-84) Avg.
Start-Up Rate	4.9	4.8	4.5	4.1	4.8	4.5	4.6
Exit Rate	4.7	4.6	4.2	4.8	4.7	4.5	4.6

1. The design effect is assumed to be 1.8 for individual pairs of months, 2.6 for half year averages, and 3.4 for the 12 month averages.

Sources: Coder (1986a), U.S. Department of Health and Human Services (1983, 1984)

Table 15. Start-Up Rates for SSI Participation (Percentages)

SIPP 84 Panel-Reference Month i to i+1 Across All Four Rotations

	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	1-5 Avg.	5-9 Avg.	1-9 Avg.
Start-Up Rate	1.4	1.2	.9	5.5	1.4	1.6	1.3	6.8	2.3	2.8	2.6
Standard Error	.8	.7	.6	1.4	.7	.8	.7	1.5	.5	.5	.4

SSI Calendar Month i to i+1

	6-7	7-8	8-9	9-10	10-11	11-12	12-13	13-14	14-15	15-16	16-17	7-12[2] Avg.	10-17[2] Avg.	7-17[2] Avg.
Start-Up Rate	.8	.6	1.1	1.0	.9	.9	.9	1.0	1.0	1.4	1.1	.9	1.0	1.0

---

[1]The design effect is assumed to be 2.6.

[2]Months 7-12 correspond to July through December of 1983 and months 13-17 correspond to January through May of 1984.

Sources: Coder (1986a), U.S. Department of Health and Human Services (1984, 1985)

Table 16. Responses for Interviews Two Through Five as a Percentage of Initially Responding Persons for 1984 SIPP Panel, NMCUES[1], and PSID.

INTERVIEW	% of Response		
	NMCUES	SIPP	PSID
(Base)	(16902)	(25138)	(18387)[2]
Second	99.5	94.4	86.6
Third	97.9	89.7	83.5
Fourth	97.1	85.9	81.5
Fifth	96.5	83.2	79.3

[1] Percentages for NMCUES include ineligible individuals, and are based on all persons in initially responding, reporting units.

[2] 1979b persons are described in most recent releases of the PSID data. An adjustment to this number was made to make it more compatible with the SIPP.

Sources: Cox, B. and S. Cohen (1985); Short, K. and E. McArthur (1986); Becketti, S., W. Gould, L. Lillard, and F. Welch, (1983)

Table 17. Overall Item Response Rate for CPS and SIPP 1985 Calendar Year Estimates[1]

Income Types	SIPP	CPS
Wage or Salary	76.1%	78.8%
Self-Employment Income	68.9	73.7%
Federal Supplemental Security Income	75.5%	78.8%
Social Security Income	72.7%	76.2%
Aid to Families with Dependent Children	77.1%	80.8%
Unemployment Compensation	72.6%	76.8%
Company or Union Pensions	70.8%	74.6%
Food Stamp Allotment	77.1%	83.9%
Veterans Compensation or Pensions	72.4%	76.7%

[1] Calendar Year item response rates are for estimates based on monthly averages.

Source: Maher, S. (1987a)

Table 18. Between Wave Transitions for Food Stamps

Sex x Interview Status

Sex	Interview State	Response Response	Response Nonresponse	Nonresponse Response	Nonresponse Nonresponse
Male	Self/Self	54.5 (456)	9.4 (79)	6.0 (50)	30.1 (252)
	Self/Proxy	45.7 (106)	12.5 (29)	8.6 (20)	33.2 (77)
	Proxy/Self	38.2 (76)	16.1 (32)	8.0 (16)	37.7 (75)
	Proxy/Proxy	37.7 (171)	12.4 (56)	5.7 (26)	44.2 (200)
Female	Self/Self	65.5 (2326)	6.8 (240)	5.2 (184)	22.6 (802)
	Self/Proxy	53.9 (125)	9.1 (21)	8.5 (20)	28.4 (56)
	Proxy/Self	43.1 (103)	9.2 (22)	9.2 (22)	38.4 (92)
	Proxy/Proxy	55.4 (92)	11.4 (19)	6.6 (11)	26.5 (44)

First entry in each cell is percent of total responses in row.
Second entry is number of responses in cell.

Table 19. Within Waves Transitions for Food Stamps

Sex x Interview State

Sex	Interview State	Response Response	Response Nonresponse	Nonresponse Response	Nonresponse Nonresponse
Male	Self/Self	57.3 (1782)	1.5 (47)	2.5 (77)	38.7 (1202)
	Proxy/Proxy	45.7 (939)	2.2 (46)	2.7 (56)	49.3 (1014)
Female	Self/Self	68.1 (7750)	1.7 (198)	2.1 (236)	28.0 (3189)
	Proxy/Proxy	59.8 (714)	1.7 (20)	1.3 (15)	37.3 (445)

First entry in each cell is percent of total responses in row.
Second entry is number of response in cell.

Table 20. Between Waves Transitions for Food Stamps

Race x Sex

Race	State	Response Response	Response Nonresponse	Nonresponse Response	Nonresponse Nonresponse
white	male	44.3 (547)	11.8 (146)	6.1 (75)	37.9 (468)
	female	59.7 (1560)	2.2 (205)	2.7 (163)	49.3 (684)
nonwhite	male	54.0 (262)	10.3 (50)	7.6 (37)	28.0 (136)
	female	68.9 (1086)	6.2 (97)	4.7 (74)	20.3 (320)

First entry in each cell is percent of total responses in row.
Second entry is number of responses in cell.

Table 21. Within Waves Transitions for Food Stamps

Race x Sex

Race	State	Response Response	Response Nonresponse	Nonresponse Response	Nonresponse Nonresponse
white	male	49.3 (1830)	2.0 (73)	3.1 (116)	45.6 (1695)
	female	64.2 (5031)	2.0 (154)	2.2 (172)	31.6 (2479)
nonwhite	male	61.2 (891)	1.4 (20)	1.6 (23)	35.8 (521)
	female	72.6 (3433)	1.4 (64)	1.7 (79)	24.4 (1155)

First entry in each cell is percent of total responses in row.
Second entry is number of responses in cell.

Table 22. Distributions of Transitions and Non-Transitions Between Waves

Source	Imputes Involved		Imputes Not Involved	
	Trans	Non-Trans	Trans	Non-Trans
Social Security	.039 (88)	.961 (2191)	.027 (499)	.973 (17785)
Veterans Compensation	.045 (11)	.955 (235)	.035 (67)	.965 (1854)
AFDC	.347 (43)	.653 (81)	.113 (257)	.887 (2014)
Food Stamps	.312 (74)	.688 (163)	.135 (773)	.865 (4940)
Child Support	.22 (24)	.78 (85)	.131 (296)	.869 (1966)
Private Pension	.103 (76)	.897 (663)	.049 (218)	.951 (4261)
Supplemental Security Income	.156 (23)	.844 (124)	.057 (128)	.943 (2113)
Unemployment Compensation	.459 (174)	.541 (205)	.192 (1119)	.808 (4720)

Trans = Transitions

Table 23. Distributions of Transitions and Non-Transitions Within Waves

Source	Imputes Involved		Imputes Not Involved	
	Trans	Non-Trans	Trans	Non-Trans
Social Security	.0003 (3)	.9997 (8594)	.008 (590)	.992 (75660)
Veterans Compensation	0 (0)	1.0 (711)	.004 (34)	.996 (8009)
AFDC	.003 (1)	.997 (301)	.030 (286)	.970 (9138)
Food Stamps	.007 (4)	.993 (596)	.043 (997)	.957 (22214)
Child Support	0 (0)	1.0 (335)	.052 (480)	.948 (8776)
Private Pension	0 (0)	1.0 (2130)	.01 (182)	.99 (18694)
Supplemental Security Income	0 (0)	1.0 (326)	.014 (125)	.986 (9121)
Unemployment Compensation	.232 (212)	.768 (701)	.108 (2616)	.892 (21656)

Trans = Transitions

Table 24. SIPP Transition Correlations

	WM	MS	MSP	PPEARN	FFINC	FFPOV	ESR	CAIDCO	AFDC	FOODST	WELFAR	FDSTA
marital	1	1.000	0.755	0.028	0.071	0.222	0.043	0.076	0.094	0.079	0.019	0.004
status	2	1.000	0.825	0.027	0.144	0.379	0.083	0.291	0.332	0.240	0.039	0.014
MS	3	1.000	0.891	0.023	0.139	0.350	0.086	0.310	0.330	0.205	0.009	0.028
	4	1.000	0.843	0.025	0.145	0.365	0.088	0.338	0.367	0.261	0.063	0.035
married	1	0.755	1.000	0.028	0.074	0.217	0.036	0.052	0.063	0.062	0.025	0.009
spouse	2	0.825	1.000	0.022	0.132	0.282	0.055	0.138	0.148	0.118	0.011	-0.002
present	3	0.891	1.000	0.025	0.129	0.270	0.067	0.187	0.199	0.128	0.011	0.032
MSP	4	0.843	1.000	0.018	0.142	0.294	0.066	0.225	0.253	0.181	0.065	0.043
person	1	0.028	0.028	1.000	0.291	0.032	0.414	0.005	0.010	0.007	0.006	0.017
earnings	2	0.027	0.022	1.000	0.323	0.032	0.532	0.025	0.022	0.030	0.018	0.036
PPEARN	3	0.023	0.025	1.000	0.321	0.027	0.523	0.014	0.014	0.033	0.011	0.050
	4	0.025	0.018	1.000	0.318	0.041	0.510	0.033	0.027	0.024	0.018	0.024
family	1	0.071	0.074	0.291	1.000	0.141	0.127	0.085	0.091	0.095	0.052	0.026
income	2	0.144	0.132	0.323	1.000	0.175	0.152	0.121	0.127	0.119	0.055	0.047
FFINC	3	0.139	0.129	0.321	1.000	0.167	0.163	0.084	0.120	0.089	0.042	0.039
	4	0.145	0.142	0.318	1.000	0.175	0.152	0.108	0.141	0.095	0.063	0.033
family	1	0.222	0.217	0.032	0.141	1.000	0.115	0.323	0.406	0.299	0.024	0.020
need	2	0.379	0.282	0.032	0.175	1.000	0.082	0.301	0.334	0.248	0.029	0.006
std.	3	0.350	0.270	0.027	0.167	1.000	0.082	0.303	0.313	0.190	-0.002	0.012
FFPOV	4	0.365	0.294	0.041	0.175	1.000	0.099	0.310	0.323	0.230	0.016	0.003
job	1	0.043	0.036	0.414	0.127	0.115	1.000	0.080	0.096	0.093	0.045	0.050
status	2	0.083	0.055	0.532	0.157	0.082	1.000	0.090	0.100	0.081	0.026	0.043
recode	3	0.086	0.067	0.523	0.163	0.082	1.000	0.089	0.091	0.077	0.016	0.054
ESR	4	0.088	0.066	0.510	0.152	0.099	1.000	0.102	0.107	0.079	0.031	0.036
medicaid	1	0.076	0.052	0.005	0.085	0.323	0.080	1.000	0.565	0.366	0.212	0.068
coverage	2	0.291	0.138	0.025	0.121	0.301	0.090	1.000	0.743	0.354	0.269	0.039
CAIDCO	3	0.310	0.187	0.014	0.084	0.303	0.089	1.000	0.590	0.279	0.186	0.037
	4	0.338	0.225	0.033	0.108	0.310	0.106	1.000	0.601	0.367	0.248	0.071
AFDC	1	0.094	0.063	0.010	0.091	0.406	0.096	0.565	1.000	0.408	0.214	0.053
coverage	2	0.332	0.148	0.022	0.127	0.334	0.100	0.743	1.000	0.411	0.291	0.079
AFDC	3	0.330	0.199	0.014	0.120	0.313	0.091	0.590	1.000	0.323	0.319	0.071
	4	0.367	0.252	0.027	0.141	0.323	0.107	0.601	1.000	0.424	0.321	0.091
foodstamp	1	0.079	0.062	0.007	0.095	0.299	0.093	0.366	0.408	1.000	0.080	0.365
coverage	2	0.240	0.118	0.030	0.119	0.248	0.081	0.355	0.411	1.000	0.103	0.419
FOODST	3	0.205	0.128	0.033	0.089	0.190	0.077	0.279	0.323	1.000	0.071	0.458
	4	0.261	0.181	0.024	0.095	0.230	0.079	0.367	0.424	1.000	0.093	0.372
welfare	1	0.019	0.025	0.006	0.052	0.024	0.045	0.212	0.214	0.080	1.000	0.159
income	2	0.039	0.011	0.018	0.055	0.029	0.026	0.269	0.291	0.103	1.000	0.182
WELFAR	3	0.009	0.011	0.011	0.042	-0.002	0.016	0.186	0.319	0.071	1.000	0.163
	4	0.063	0.065	0.018	0.063	0.016	0.031	0.248	0.321	0.093	1.000	0.148
foodstamp	1	0.004	0.009	0.017	0.026	0.020	0.050	0.068	0.053	0.365	0.159	1.000
allotment	2	0.014	-0.002	0.036	0.046	0.006	0.043	0.039	0.079	0.419	0.182	1.000
FDSTA	3	0.028	0.032	0.050	0.039	0.012	0.054	0.037	0.071	0.458	0.163	1.000
	4	0.035	0.043	0.024	0.033	0.003	0.036	0.071	0.091	0.372	0.148	1.000

Months 5-32, Full Panel Research File: Observations fully interviewed

Source: Young (1989)

REFERENCES

Abowd, J. and A. Zellner (1985), "Application of Adjustment Techniques to U. S. Gross Flow Data," Proceedings of the Conference on Gross Flows in Labor Force Statistics, p. 45-61.

Becketti, S., W. Gould, L. Lillard, and F. Welch, (1983) "Attrition from the PSID," Unicom Research Corp., Santa Monica, CA. November 11, 1983.

Burkhead, D. and J. Coder (1985), "Gross Changes in Income Recipiency from the SIPP," Proceedings of the Section on Social Statistics, American Statistical Association.

Carlson, S. and Dalrymple, R. (1986), "Food Stamp Participation: A Comparison of SIPP With Administrative Records," Second Annual Research Conference Proceedings, U.S. Bureau of the Census, P. 208-234.

Coder, J. (1986a), "Monthly Transitions from the SIPP Longitudinal Research File," Internal Census Bureau memorandum from Coder to Schneider, May 20, 1986.

Coder, J. (1986b), "Some Weighted Estimates from the 1983-84 SIPP Longitudinal Research File," Internal Census Bureau memorandum from Coder to Green, October 8, 1986.

Coder, J. (1987a), "Comparison of 1984 and 1985 SIPP Panel Estimates," Internal Census Bureau memorandum from Coder to Green, February 27, 1987.

Coder, J. (1987b), "Some Notes on Quality of SIPP Unemployment Compensation Data," Internal Census Bureau Memorandum from Coder to Green, April 7, 1987.

Coder, J. (1987c), "Recommended Changes to the Core of the 1988 SIPP Panel Questionnaires," Internal Census Bureau Memorandum from Coder to Green, April 20, 1987.

Coder, J. (1987d), "Wage and Salary Data from the Annual Roundup for 1984," Internal Census Bureau memorandum from Coder to Green, April 20, 1987.

Coder, J. (1987e), " Additional Tabulations of Wave 6 Tax Data," Internal Census Bureau memorandum from Coder to Green, October 29, 1987.

Coder, J., Burkhead D., Feldman-Harkins, and J. McNeil (1987) "Preliminary Data From the 1983-84 Longitudinal Research File," SIPP Working Paper Series No. 8702, Bureau of the Census.

Cox, B. and S. Cohen, (1985) "Methodology Issues for Health Care Surveys." Published by Marcel Dekker, Inc.

Curtin, R.T., F.T. Juster and J.N. Morgan (1987), "Survey Estimates of Wealth: An Assessment of Quality," presented at the conference of Research in Income and Wealth, Baltimore, Maryland, March 27-28.

Dahmann, D. and E. McArthur (1987), "The Analysis of Geographical Mobility and Life Events with the SIPP," SIPP Working Paper Series No. 8720, Bureau of the Census

Fuller, W. and T. Chin Chua (1985), "Gross Change Estimation in the Presence of Response Error," Proceedings of the Conference on Gross Flows in Labor Force Statistics, p. 65-77.

Fuller, W. and T. Chin Chua (1986), "Response Error Models for Changes in Multinomial Variables," Second Annual Research Conference Proceedings, U.S. Bureau of the Census, p. 425-441.

Hill, M. (1987), "Marital Status and Changes therein in SIPP," Internal Census Bureau draft memorandum from Hill to Census staff, February 20, 1987.

Hock, E. (1988), "SIPP 1984: Specifications for Panel File Longitudinal Weighting of Persons," Internal Census Bureau memorandum from Waite to Walsh, December 1988.

Holt, M. (1986) "Evaluation of Training Materials and Methods for the SIPP," SIPP Working Paper Series No. 8608, Bureau of the Census.

Hubble, D. L. and D. R. Judkins (1986), "Measuring the Bias in Gross Flows in the Presence of Auto-Correlated Response Errors," Proceedings of the Section on Survey Research Methods, American Statistical Association,

Huggins, V. (1987), "SIPP: Evaluation of Missing Wave Data," Internal Census Bureau memorandum from Singh to Shapiro, February, 1987.

Judkins, D. R. (1985), "SIPP -- Idea for Estimation of Response Variance Without Reinterview," Internal Census Bureau Memorandum from Judkins for documentation, February 25, 1985.

Judkins, D. R. (1986), "SIPP Gross Flows: Validation of Food Stamp Turnover," Internal Census Bureau memorandum from Judkins to Singh, October 16, 1986.

Kalton, G. (1986), "Handling Wave Nonresponse in Panel Surveys," Journal of Official Statistics, Vol. 2, No. 3, 1986, pp. 303-314.

Kalton, G. and M.E. Miller (1986), "Effects of Adjustments for Wave Nonresponse on Panel Survey Estimates," Proceedings of the Section on Survey Research Methods, American Statistical Association.

Kim, J. (1985), "SIPP: Evaluation of Selected SIPP Estimates," Internal Census Bureau memorandum from Singh to Shapiro, December 9, 1985.

King, K. (1985), "SIPP 85: Cross-Sectional Weighting Specifications for Wave 1 -- Revision," Internal Census Bureau memorandum from Jones to Walsh, November 21, 1985.

King, K. and J. Kim (1986), "SIPP: Cross-Sectional Weighting Specifications for the Second and Subsequent Waves," Internal Census Bureau memorandum from Jones to Walsh, June 19, 1986.

Kobilarcik, E., C. Alexander, R. Singh and G. Shapiro (1983), "Effects of Reference Period on Reporting of Victimization," Proceedings of the Section on Survey Research Methods, American Statistical Association.

Kobilarcik, E. and R. Singh (1986), "SIPP Longitudinal Estimation for Persons' Characteristics," Proceedings of the Section on Survey Research Methods, American Statistical Association.

Kominski, R. (1987), "Estimate of Educational Financing Recipients from SIPP Wave 6, 84 Panel," Internal Census Bureau memorandum from Kominski to Norton, June 1, 1987.

Lamas, E. and J. McNeil (1987), "An Analysis of the SIPP Asset and Liability Feedback Experiment," SIPP Working Paper Series No. 8725, Bureau of the Census.

Linebarger, J. (1986), "Survey of Income and Program Participation: Verification Results of the August Data Keying," Internal Census Bureau memorandum from Linebarger to Bowie, September 10, 1986.

Maher, S. (1987a), "SIPP: Overall Item Response Rates for 1985 Calendar Year Estimates," Internal Census Bureau memorandum from Maher for documentation, July, 1987.

Maher, S. (1987b), "SIPP 84: Validation of AFDC Turnover," Internal Census Bureau memorandum from Singh to Shapiro, August 14, 1984.

Maher, S. (1987c), " SIPP 84: Validation of SSI Turnover," Internal Census Bureau memorandum from Singh to Shapiro, October 30, 1987.

Matchett, S.D. (1987), "Debriefing SIPP Respondents," Internal Census Bureau memorandum from Matchett to All Regional Directors, June 4, 1987.

McArthur, E. and K. Short (1986), "Measurement of Attrition from the SIPP through the Fifth Wave of 1984 Panel," Internal Census Bureau memorandum from McArthur and Short, April 10, 1986.

Moore, J. (1986), "Background Paper for the Meeting of the ASA/SSRC Joint Working Group on the Technical Aspects of SIPP," U.S. Bureau of the Census, Washington, D.C., October 1986.

Moore, J. and D. Kasprzyk (1984), "Month-to-Month Recipiency Turnover in the ISDP," Proceedings of the Section on Survey Research Methods, American Statistical Association.

Nelson, D., D. McMillen, and D. Kasprzyk (1985), "An Overview of the Survey of Income and Program Participation," SIPP Working Paper Series, No. 8401, Update, U.S. Bureau of the Census, Washington, D.C., December 1985.

O'Muircheartaigh, C. (1986), "Correlates of Reinterview Response Inconsistency in the Current Population Survey (CPS)," Second Annual Research Conference Proceedings, U.S. Bureau of the Census, p. 208-234.

Petroni, R. (1986), "SIPP: Recall Effect Study," Internal Census Bureau memorandum from Singh to Shapiro, October 2, 1986.

Poterba, J. and L. Summers (1985), "Adjusting the Gross Changes Data: Implications for Labor Market Dynamics," Proceedings of the Conference on Gross Flows in Labor Force Statistics, p. 81-95.

Ruggles, P. and William, R. (1986), "Transitions In and Out of Poverty: New Data from the Survey of Income and Program Participation," presented at the Annual Meetings of the Allied Social Science Associations, December, 1986.

Ryscavage, P. and A. Feldman-Harkins (1988), "A Comparison of Gross Changes in Labor Force Status from SIPP and CPS," Presented at the Conference on Individuals and Families in Transition: Understanding Change Through Longitudinal Data, Annapolis, Maryland, March 16-18, 1988.

Schwanz, D. (1986), "Formation of the AHS Response Variance/Bias Research Project Work Group," Internal Census Bureau Memorandum to J. Cannon, J. Linebarger, E. Montfort, J. Paletta, and B. Phalen, May 13, 1986.

Short, K. and E. McArthur (1986), "Life Events and Sample Attrition in the Survey of Income and Program Participation," Proceedings of the Section on Social Statistics, American Statistical Association.

Singh, R. (1986), "SIPP High Priority Projects," Internal Census Bureau memorandum from Singh to Kasprzyk, August 7, 1986.

U.S. Bureau of the Census "The Current Population Survey Reinterview Program, Jan. 1961 through Dec. 1966," Technical Paper #19, Washington, D.C.: Government Printing Office, Washington, D.C. 1968.

U.S. Bureau of the Census, Current Population Reports, Series P-70, Nos. 1-5, U.S. Government Printing Office, Washington, D.C. 1984-1985.

U.S. Bureau of the Census, Current Population Reports, Series P-70, No. 7, "Household Wealth and Asset Ownership: 1984," U.S. Government Printing Office, Washington, D.C. (1986).

U.S. Bureau of the Census, Statistical Abstract of the United States: 1986, (106th edition) Washington, D.C. 1986.

U.S. Bureau of the Census, Current Population Reports, Series P-70, No. 9, "Who's Minding the Kids? Child Care Arrangement: Winter 1984-85," U.S. Government Printing Office, Washington, D.C. 1987.

U.S. Bureau of the Census and U.S. Bureau of Labor Statistics, Proceedings of the Conference on Gross Flows in Labor Force Statistics, Washington, D.C. 1985.

U.S. Department of Health and Human Services, "Quarterly Public Assistance Statistics," July-Sept., Oct.-Dec., 1983; Jan.-Mar., Apr.-June 1984, Washington, D.C.

U.S. Department of Health and Human Services, "Social Security Bulletin," Volume 47, No. 3, 1984 and Volume 48, No. 4, 1985.

Urban Institute, "Effects of Legislative Changes in 1981 and 1982 on the Food Stamp Program," Volume II, Appendix E, Washington, D.C. 1985.

Valentine, C.A. and B. Valentine (1971), "Missing Men - A Comparative Methodological Study of Underenumeration and Related Problems." Unpublished report.

Vaughan, D.R. (1988), "Reflections on the Income Estimates from the Initial Panel of the Survey of Income and Program Participation (SIPP)," presented at the SIPP conference on Individuals and Families in Transition: Understanding Change through Longitudinal Data, Annapolis, Maryland March 16-18, 1988.

Weidman, L. (1986), "Investigation of Gross Changes in Income Recipiency from the Survey of Income and Program Participation," Proceedings of the Section on Survey Research Methods, American Statistical Association.

Weidman, L. (1987), "Investigation of Possible Causes of Transition Patterns from SIPP," Internal Census Bureau Report, February 1987.

Young, N. (1989), "Wave Seam Effects in the SIPP -- Implications for Analysis," Forthcoming under SIPP Working Paper Series.

THE SURVEY OF INCOME AND
PROGRAM PARTICIPATION

THE SURVEY OF INCOME AND PROGRAM PARTICIPATION: AN OVERVIEW AND DISCUSSION OF RESEARCH ISSUES

No. 8830 77

Daniel Kasprzyk, Census Bureau

December 1988

U.S. Department of Commerce BUREAU OF THE CENSUS

This paper reports the general results of the research undertaken by Census Bureau staff and was prepared at the request of staff from Statistics Sweden. The views expressed are attributable to the author and do not necessarily reflect those of the Census Bureau. This paper brings together new material with that found in "An Overview of the Survey of Income and Program Participation: Update 1," by Dawn Nelson, David B. McMillen, and Daniel Kasprzyk, SIPP Working Paper Series No. 8401, and "Research Issues in the Survey of Income and Program Participation," Daniel Kasprzyk, Survey Methodology (1988).

1. Introduction

The Survey of Income and Program Participation (SIPP) is an on-going nationally representative household survey program of the U.S. Bureau of the Census. It provides comprehensive information on the economic resources of the American people and on how public transfer and tax programs affect their financial circumstances. The data from the SIPP provide government policymakers with an information base for studying government tax and transfer programs, for estimating future program costs and coverage, and for assessing the effects of proposed policy changes. The SIPP is designed to improve the measurement of information related to the economic situation of households and persons in the United States, and is the culmination of a large-scale development program, the Income Survey Development Program (ISDP), which examined concepts, procedures, questionnaires, and recall periods (Ycas and Lininger, 1981).

The need for a survey like SIPP arose because of the limitations of the March Income Supplement of the Current Population Survey (CPS), the principal source of information on the distribution of household and personal income in the United States. These limitations are inherent in the survey design, survey instrument, and survey procedures and can not be easily modified. As a consequence the Income Survey Development Program was established in 1975 by the U.S. Department of Health and Human Services to develop methods to overcome the principal shortcomings of the CPS--1) the underreporting of property income and other irregular sources of income; 2) the underreporting and misclassification of participation in major income security programs and other types of information that people generally find difficult to report accurately (for example, monthly detail on income earned during the year); and 3) the lack of information necessary to analyze program participation and eligibility. Several features distinguish the field tests of the ISDP from other data collections, particularly the CPS. They include: 1) interviews were obtained at regular intervals within a year; 2) most types of income were reported on a monthly basis; 3) income was reported on an individual basis; 4) individuals were followed over the survey period to obtain data on changes in income and family composition; and 5) information was collected on special topics such as disability, child care, fertility, net worth, and taxes paid to provide insight into the context of program benefits, program dependency, and overall economic well-being. Because the ISDP was the predecessor to SIPP, many characteristics of the ISDP can be seen in the SIPP, including the survey design, content, and questionnaire format.

This paper provides basic background information on the survey design and content as a prelude to its more specific goal of reviewing specific methodological, survey design, and statistical issues of concern to the program, including (1) questionnaire design; (2) data collection, including respondent rules, data collection mode, length of reference period, and rules for following movers; (3) concepts, design, and estimation; and (4) response error.

2. What is the SIPP?

The SIPP is a continuous household survey program of the U.S. Bureau of the Census with interviews of sample members conducted every 4 months for 32 months. Since a new sample is introduced each year, it may be thought of as a rotating

panel survey. Its principal features, about which more will be said below, are

a. most income is reported on a monthly basis;
b. individuals are followed for changes in income and household composition;
c. income is reported on an individual basis;
d. information on special topics is linked to the income data.

The purpose of the SIPP is to provide data to understand more completely the economic well-being of the Nation through (a) better measurement of income and program participation and (b) an expansion in what is meant by economic well-being--assets, liabilities, employer-provided benefits, and demographic and historical data.

2.1 Design Features

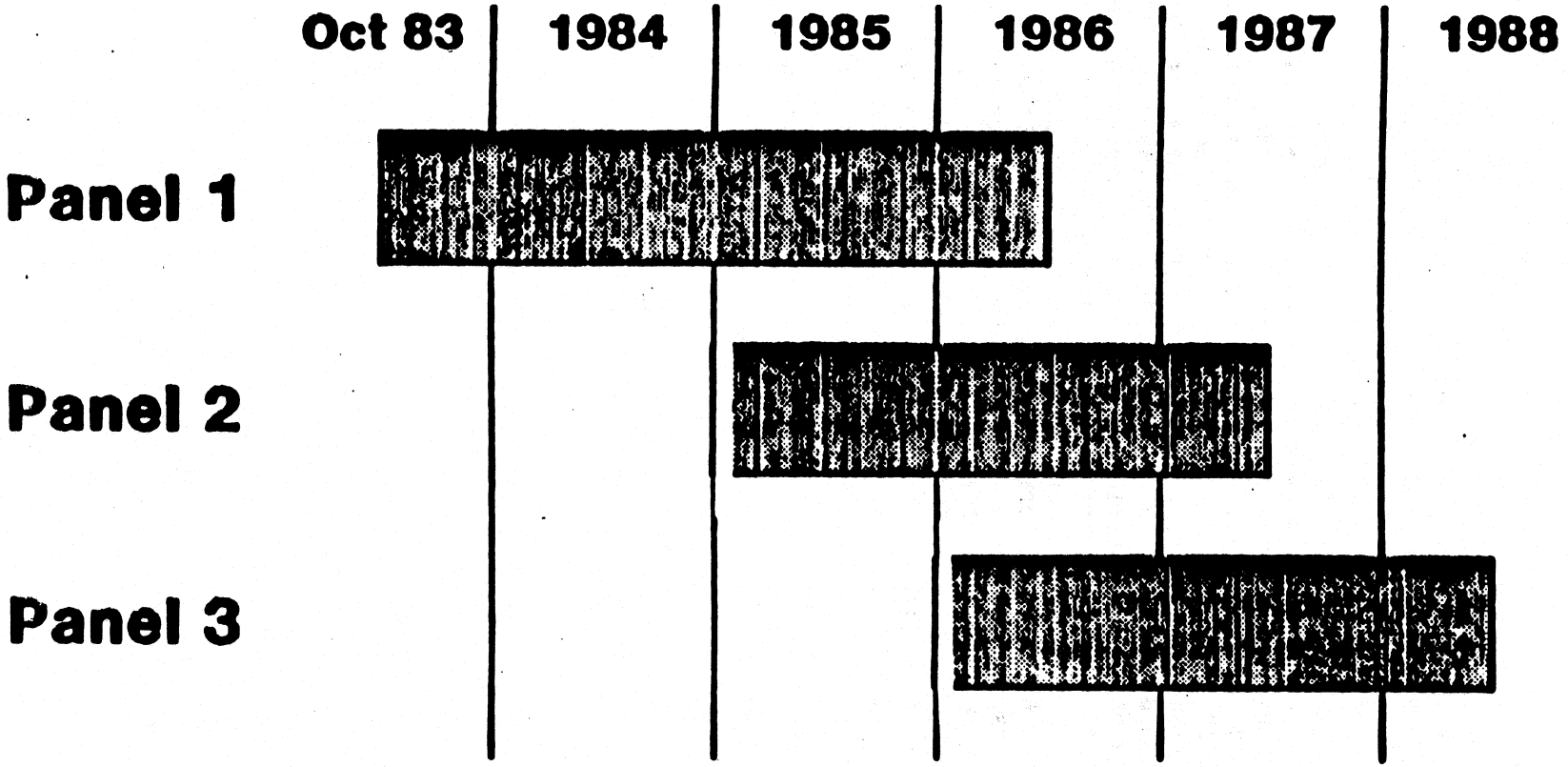
The primary goals in designing SIPP were to improve reporting of income and other program-related data and to do it in a way that would allow the analysis of changes over time at a microlevel. The design also had to accommodate the collection of a large quantity of information in a flexible manner that allowed some information to be collected more frequently than other information. These goals were met principally by using a survey design in which the same people are interviewed more than once. Persons at households selected for a sample panel are interviewed about their income and other topics once every 4 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, or vice versa, are also interviewed. In this way, a highly detailed record is built up over time for each person and household in a sample panel. This design minimizes the need for sample persons to recall most of the information for longer than a few months and reduces the number of questions asked in one interview.

To further enhance the estimates of change, particularly year-to-year change, a new sample panel is introduced every year instead of at the conclusion of a panel. Consequently, two or sometimes three panels are in the field concurrently, as is illustrated in figure 1. This overlapping panel design allows cross-sectional estimates to be produced from a larger, combined sample that is about double in size when two panels overlap.

The first SIPP panel, designated as the 1984 Panel but implemented in October 1983, started with approximately 20,000 interviewed households. The second panel, i.e., the 1985 Panel, began in February 1985 with arour 14,000 interviewed households. Because of budget constraints, new panels about 12,000 interviewed households are now fielded every February.

The reference period for the primary survey items is the 4 month re q the interview; for example, in February, the reference period is pr n October through January. When the household is interviewed again Jur h reference period is February through May. To create manageable in vview and processing work loads each month instead of one large work loa every 4 months, the sample households within a given panel are divided o fo subsamples of nearly equal size. These subsamples are called rot on g ps, and one rotation group or one-fourth of the sample is interviewed ach mon h.

Figure 1.
Overlapping Panels



Thus, it takes 4 consecutive months to interview the entire sample. This 4-month period of interviewing is called a wave. The following is an illustration of the relationship between waves, rotation groups, interview months, and reference periods in the 1987 Panel (Note minor but important differences in these relationships exist in the 1984, 1985 and 1986 Panels). The basic relationships are the same in subsequent panels. Looking at Wave 1 in figure 2, persons interviewed in Febr[illegible]y report data for the period October through January; in March another rota[illegible]n group report for November through February, and so forth for each of the fou[illegible] rotation groups. Notice that each rotation group within a wave uses a different reference period, namely, the 4 months preceding the interview month. As a result, data are available for 7 months at the conclusion of Wave 1 interviewing although each month is not represented by the full sample. For example, the October and April data will only be available for one rotation group (rotation groups 2 and 1, respectively); the November and March data for two rotation groups (rotation groups 2 and 3, and 4 and 1, respectively); and the December and February data for three rotation groups. Only the January data will be represented by the full sample (rotation groups 1-4). In Wave 2, the persons originally interviewed in February are interviewed again in June for information on the months of February through May. In July, the March respondents are asked to report data for March through June, and so on. After two interviews with the same rotation group, eight consecutive months of data are available. Then, if data collected in Wave 1 are used together with Wave 2 data, estimates for February through April can be produced using all four rotation groups. Thus, to produce calendar quarter estimates for the full sample, it is necessary to work with more than one wave of data. In the same way, data corresponding to a calendar year can only be obtained by matching data from four consecutive waves of interviewing; e.g., data collected in interviews conducted from February 1987 (part of Wave 1) through April 1988 (part of Wave 4) can be merged to produce monthly data covering calendar year 1987.

2.2 Survey Content

Each interview is planned to take about 30 minutes of a respondent's time and includes content that is divided into three main groups of questions. The substance of two of these groups should be essentially the same for each wave and for each panel. The third group of questions covers topics that will change in each wave of a panel. This will allow for the inc[illegible]sion of some new content in each panel, although many of the topics will be r[illegible]eated across all the panels. Each rotation group in a wave is administered the same set of questions although the reference period is different as explained above.

The first group of questions are control card items. The control card is a separate document from the questionnaire and serves several important funct[illegible]ns. The control card is used to list every person residing at an address and to record basic social and demographic characteristics (age, race, sex, and so forth) for each person at the time of the initial interview. The card is reused at subsequent interviews to record changes in characteristics such as age, education attainment, and marital status; and to record the dates when persons enter or leave the household. Finally, during each interview, information on each source of income received and the name of each job or business is transcribed to the card so that this information can be used in the updating process at the next interview.

Figure 2. Relationship Between SIPP Interview Months and Reference Periods: 1987 Panel

Wave	Rotation Group	Interview Month	Reference Periods: 1st Qtr Data	2nd Qtr Data	3rd Qtr Data	4th Qtr Data	
1	2	FEB OCT NOV DEC	JAN				
	3	MAR NOV DEC	JAN FEB				
	4	APR DEC	JAN FEB MAR				
	1	MAY	JAN FEB MAR	APR			
2	2	JUN	FEB MAR	APR MAY			
	3	JUL	MAR	APR MAY JUN			
	4	AUG		APR MAY JUN	JUL		
	1	SEP		MAY JUN	JUL AUG		
3	2	OCT		JUN	JUL AUG SEP		
	3	NOV			JUL AUG SEP	OCT	
	4	DEC			AUG SEP	OCT NOV	
	1	JAN			SEP	OCT NOV DEC	
4	2	FEB				OCT NOV DEC	JAN
	3	MAR				NOV DEC	JAN FEB
	4	APR				DEC	JAN FEB MAR
	1	MAY					JAN FEB MAR APR
5							
6		Continues in the pattern established above.					
7							
8							

The second major group of questions form the core portion of the questionnaire, which is divided into five sections. The core set of questions is asked at the first interview and then updated in each subsequent interview. The first section of the core collects the basic labor force participation data for the 4 reference months. In addition, this first section of the core collects much of the information on the receipt of income from various sources during the 4-month reference period. This includes income from government sources such as Aid to Families with Dependent Children, Supplemental Security Income, General Assistance, and Workmen's Compensation. Respondents are also asked about Social Security and other retirement income. The receipt of miscellaneous sources of income such as alimony, child support, interest from savings, income for foster child care, and educational assistance is also identified. In addition, questions on major sources of noncash benefits such as food stamps, Medicaid, Medicare, and health insurance coverage are included in this section.

The second section of the SIPP core questionnaire collects information associated with wage and salary earnings. This section includes information on industry and occupation as well as hourly earnings for up to two jobs. Data are collected for two jobs held either concurrently or sequentially during the 4-month reference period.

The third section of the core collects data on self-employment earnings and specific information about the kind of self-employment--whether it was incorporated, sole proprietorship, or partnership--and the profits and losses from the business.

The fourth section is identified as the general amounts section. This section of the questionnaire collects monthly amounts received from the income sources identified in the first section. Space is provided for amounts from up to six income sources.

The fifth and last section of the core questionnaire collects amounts of income earned from asset holdings. Asset sources include savings accounts, bonds, stocks, and rental property, as well as others. Information is collected for the 4-month reference period on both individual and joint recipiency.

The third major question group consists of the various supplements or topical modules that are included in waves following the initial interview. The administration of a module is possible in Waves 2 through 8 because less time is required to update the core information after the first interview. The topical modules cover areas that do not require examination every 4 months and may use a different reference period than the core questions. The modules provide a broader context for analysis by obtaining information on a variety of topics not covered in the core portion of the questionnaire. The module data may be analyzed independently or in conjunction with the control card items or core data. Frequently, a module is administered at the same time in concurrent panels so that the data may be combined to improve the reliability of the analyses.

There are two types of topical modules: fixed and variable. The fixed topical modules are designed to be conducted on a regular basis to augment the core data. They are considered necessary to meet the survey's goals and objectives. Although the topics are "fixed," the questions in these modules may be modified from time-to-time to accommodate conceptual changes or to make improvements in

collecting these data. An example of a fixed topical module is the annual "round-up" module on earnings and benefits. This module obtains wages and salary data from W-2 forms (a wage and tax statement filed by each employer for each employee) and estimates of annual self-employment for each appropriate person in the fifth and eighth interviews in each panel. Another fixed module administered at the same time obtains property income and tax-related information; e.g., filing status and taxes paid, to allow the estimation of tax incidence, disposable income, and the simulation of tax policy alternatives.

The variable topical modules are designed to satisfy the special programmatic needs of other Federal agencies. Time is set aside for variable modules to meet special content needs that develop as the survey continues. An example of a variable topical module is the child care topical module administered in the 1984 Panel. Variable topical modules may be repeated in subsequent waves or panels as necessary. Figure 3 contains a list of the fixed and variable modules scheduled for the 1984-1987 Panels.

2.3 Operational Procedures

Data collection operations are managed through the Census Bureau's 12 permanent regional offices. A staff of interviewers assigned to SIPP conduct interviews by personal visit each month. Self-response is required for each person 15 years old and older who is present at the time of interview and is obtained in about 65 percent of the cases each wave. A proxy respondent is asked to provide information for those who are not available. Telephone interviewing occurs in about 5 percent of the cases to obtain missing information, to interview persons who will not or cannot participate otherwise, or to interview persons who have moved far outside the interviewing area. Most of the interviewing is completed during the first 2 weeks of a month.

For cost reasons, personal visit interviews are only conducted at new addresses that are within 100 miles of a SIPP sampling area; telephone interviews are used otherwise. Persons who move into an institution, Armed Forces barracks, or outside the United States are not interviewed at the new location. When a sample person leaves an institution, interviewing resumes. (This procedure, however, was not implemented until the spring of 1985.)

When an original sample person (those interviewed in the first wave) moves in with other people, all of the additional persons (age 15 or older) are interviewed in subsequent waves. Additional persons (age 15 or older) who move in with original sample persons are interviewed also. These additional persons are considered part of the sample and are interviewed only while residing with the original sample person(s). These provisions were adopted because most types of analysis using SIPP data will focus on the household and family situation of individuals. (See papers by Kalton and Lepkowski (1985) and Jean and McArthur (1984) for further discussion of following movers.)

3. Questionnaire Design

The preceding section briefly described SIPP design, content, and operational features. It serves as background information to the discussion of research issues in the SIPP. The first topic--questionnaire design has been and will continue to be an important issue in the SIPP. Investigations have been conducted concerning: a) format of the questionnaire; b) independent versus

Figure 3. SURVEY OF INCOME AND PROGRAM PARTICIPATION
TOPICAL MODULE SCHEDULE
1984/1985 Panels

INTERVIEW DATES	1984 PANEL Wave	Fixed Topical Module	Variable Topical Module	1985 PANEL Wave	Fixed Topical Module	Variable Topical Module
Oct 83-Jan 84	1	None	None			
Feb 84 Apr 84	2	None	None			
May 84 Aug 84	3	Health and Disability Work History Education History				
Sept 84 Dec 84	4	Assets Liabilities	Pension Plan Coverage Characteristics of Job from which Retired Retirement Plans and Expectations Housing Costs and Conditions Energy Usage			
Jan 85 Apr 85	5		Child Care Arrangements and Expenses Welfare History Child Support Support for Non-Household Members Reasons for Not Working Reservation Wage Work Related Expenses	1	None (Feb 1985-May 1985)	None
May 85 Aug 85	6	Annual Income Employee Benefits Educational Financing and Enrollment	Training Questions (ETA)	2	None (June 1985-[illegible])	
Sep 85 Dec 85	7	Assets Liabilities	Pension Plan Coverage Update	3	Assets Liabilities	
Jan 86 Apr 86	8	Marital History Fertility History Migration History	Household Relationships Support for Non-Household Members Work Related Expenses	4	Marital History Fertility History Migration History	Household Relationships Support for Non-Household Members Work Related Expenses

Figure 3. SURVEY OF INCOME AND PROGRAM PARTICIPATION
TOPICAL MODULE SCHEDULE
1984/1985 Panels Continued

INTERVIEW DATES	1984 PANEL			1985 PANEL		
	Wave	Fixed Topical Module	Variable Topical Module	Wave	Fixed Topical Module	Variable Topical Module
May 86 Aug 86	9	Annual Income Taxes Individual Retirement Accounts Educational Financing And Enrollment		5	Annual Income Taxes Individual Retirement Accounts Educational Financing And Enrollment	
Sept 86 Dec 86				6		Child Care Arrangements Child Support Agreements Support for Nonhousehold Members Job offers Health Status and Utilization of Health Care Services Long-Term Care Disability Status of Children
Jan 87 Apr 87				7	Assets Liabilities	Pension Plan Coverage Lump Sum Distributions from Pension Plans Characteristics of Job from which Retired Characteristics of Home Financing Arrangements
May 87 Aug 87				8	Annual Income Taxes Individual Retirement Accounts Educational Financing and Enrollment	

Figure 3. SURVEY OF INCOME AND PROGRAM PARTICIPATION
TOPICAL MODULE SCHEDULE
1986 Panel Continued

INTERVIEW DATES	Wave	1986 PANEL Fixed Topical Module	Variable Topical Module
Feb 86 May 86	1	None	None
June 86 Sept 86*	2	Fertility History Marital History Migration History Recipiency History Employment History Work Disability History Education and Trainings History Household Relationships	Personal History
Oct 86 Dec 86	3		Child Care Arrangements Child Support Agreements Support for Nonhousehold Members Job Offers Health Status and Utilization of Health Care Services Long-Term Care Disability Status of Children
Jan 87 Apr 87	4	Assets Liabilities	Pension Plan Coverage Lump Sum Distributions from Pension Plans Characteristics of Job from which Retired Characteristics of Home Financing Arrangements

* These modules are collectively identified as the Personal History Topical Module.

Figure 3. SURVEY OF INCOME AND PROGRAM PARTICIPATION
TOPICAL MODULE SCHEDULE
1986 Panel Continued

INTERVIEW DATES	Wave	1986 PANEL Fixed Topical Module	Variable Topical Module
May 87 Aug 87	5	Annual Income Taxes Individual Retirement Accounts Educational Financing And Enrollment	
Sept 87 Dec 87	6		Child Care Arrangements Child Support Agreements Support for Nonhousehold Members Work Related Expenses Housing Costs Energy Usage
Jan 88 Apr 88	7	Assets Liabilities	

Note the 1986 Panel did not contain an eighth interview due to budget reasons.

Figure 3. SURVEY OF INCOME AND PROGRAM PARTICIPATION
TOPICAL MODULE SCHEDULE
1987 Panel Continued

INTERVIEW DATES	Wave	1987 PANEL Fixed Topical Module	Variable Topical Module
Feb 87 May 87	1	None	
June 87 Sept 87	2	Fertility History Marital History Migration History Recipiency History Employment History Work Disability History Education and Training History Family Background Household Relationships	Personal History
Oct 87 Jan 88	3		Child Care Arrangements/ Child Support Agreements Support for Non-Household Members Work-Related Expenses Shelter Costs/Energy Usage
Feb 88 May 88	4	Assets and Liabilities Real Estate Property and Vehicles	
June 88 Sept 88	5	Annual Income and Retirement Accounts Taxes School Enrollment and Financing	

Figure 3. SURVEY OF INCOME AND PROGRAM PARTICIPATION
TOPICAL MODULE SCHEDULE
1987 Panel Continued

INTERVIEW DATES	Wave	1987 PANEL Fixed Topical Module	Variable Topical Module
Oct 88 Jan 89	6		Work Schedule Child Care Arrangements Child Support Agreements Support for Non-Household Members Health Status and Utilization of Health Care Services Long-Term Care Disability Status of Children
Feb 89 May 89	7	Assets and Liabilities Real Estate Property and Vehicles Work Disability Assets Deductible expenses for shelter, medical care, dependent care	Eligibility
June 89 Sept 89	8	Annual Income and Retirement Accounts Taxes School Enrollment and Financing	

dependent updates of income sources; c) the use of a section of the questionnaire to obtain data missing from an earlier wave; d) the use of data in the "annual round-up" to help develop calendar-year income estimates, and e) approaches to the collection of employer-provided benefits.

3.1 Questionnaire Format

The principal effort of the ISDP was directed to overcoming problems which resulted in underreporting and misclassification of income in the CPS March Supplement. In an ISDP field test, two questionnaire approaches were developed. For simplicity, one version may be referred to as the "short" form and the other as the "long" form.

The short-form approach attempted to gather income data directly while keeping respondent burden at a moderately low level. For each household member, questions were asked directly about the receipt of certain income types. If income were received, the amount received during the reference period was determined before proceeding to the next source of income.

The general strategy of the long-form approach was to isolate events, experiences, and other attributes associated with the receipt of specific types of income. This form contained an extensive set of probes about the receipt of income and lengthy questions to ascertain income amounts. Amounts associated with specific income types were not obtained until all sources of income were determined.

The hypothesis tested was that the long-form approach produces more complete and accurate reporting of income; Olson (1980) provides a summary of the analysis conducted on the two questionnaire formats. Several approaches to the analysis were implemented and are discussed in Olson's summary: (1) staff observation of training and interviewing; (2) debriefing sessions of interviewers and observers; (3) case-by-case reviews of completed questionnaires; (4) analysis of survey and item response rates; and (5) data analyses focussing on the quality of the data collected and questionnaire edit failures, especially those associated with the inability of the interviewer to follow questionnaire skip patterns. The form adopted for further research and ultimately the SIPP was a variation of the long form. The long form was perceived by both interviewers and respondents as less burdensome and also was shown to have higher income-reporting rates.

Another experiment with questionnaire formats was also included in the ISDP; this experiment contrasted a household-screening format with a person-based approach which was based on a revised version of the questionnaire used in the April 1978 CPS Income Supplement Test. The latter version was intended to reduce burden by asking a single household respondent whether anyone in the household received a particular kind of income during the reference period. Each affirmative response was followed by a question to identify exactly which household member(s) received that type of income. Complete recipiency for all household members was recorded before asking about amounts of income received by specific individuals. This approach was expected to reduce interview time without reducing data quality.

The approach above was contrasted with a person-based approach. Under this approach, questions on all sources of income were asked of the first household

member, then repeated for the second, and so on. A separate form was filled out for each adult in a sample household, but extensive use was made of skip instructions and check items to reduce the number of questions asked of any one respondent.

Differences in the quality of the data obtained with the two questionnaire formats and differences in the interview times appeared slight. Large differences were not observed between the two approaches in estimates of income-recipiency rates, and in the incidence of "don't know" and "refusals." Interview time, expected to be significantly less under the household questionnaire approach, was about 5 minutes less per household and about 3 minutes less per person than the person approach. Since the household-screening format did not offer a significant improvement over the person-based approach, this person based format, with modest improvements and refinements, was adopted for SIPP.

3.2 Independent versus Dependent Updates

Questionnaire design issues and discussions concerning data collection procedures continue to be a part of the SIPP program. The general issue is whether interviews conducted without the use of responses from previous interviews (the so-called independent approach) produce better estimates than interviews conducted using the previous interview responses to remind respondents of earlier statuses (the so-called dependent-interview approach). In the SIPP a dependent approach is used to update income receipt patterns at each interview.

Figure 4 exhibits the questions designed to update sources of income. In order to conduct the interview, the interviewer must transcribe income sources reported at the previous interview from the control card onto the questionnaire. The approach has not been systematically evaluated, but it is apparent from several analyses that the approach is not working as well as some had expected.

Data problems related to the correct timing of changes in income sources exist. As will be discussed in section 6, there appear to be too many transitions in receipt of income sources between the first month of the previous interview and the last month of the current interview.

A similar dependent approach to data collection is also possible with the data collected in the SIPP on personal net worth. These data are obtained at two points in time, one year apart. Specifically, data on asset and liability values, collected in Wave 4 of the 1984 Panel, were provided to one-half of the respondents interviewed in the Wave 7 interview. To examine differences between the dependent and independent approach, one-half the sample in Wave 7 was provided information collected on asset and liability values collected in Wave 4, while the other half was not provided the previously reported information.

The rationale for this dependent or "feedback" approach was that respondents would provide more accurate estimates of change if they were first reminded of the amount they reported the previous year. If respondents know the amount of the change in asset values and were reminded of their beginning balance, then presumably their reporting of the current balance would be consistent with the true amount of change over the period. Lamas and McNeil (1987) analyze these data, but give no definite answer about the impact of the feedback approach since benchmark data are not available. They do, however, say that the dependent interview did not affect cross-sectional estimates and that the approach

Figure 4

Section 1 — LABOR FORCE AND RECIPIENCY (Continued)

11a. According to the information we obtained last time, . . . had received *(Read income types in 11b, column (2))* during *(8 months ago)* through *(5 months ago)*. Was this information recorded correctly?

1251 1 ☐ Yes
2 ☐ No — *Resolve problems and make appropriate entries in 11b, column (5)* } *Ask 11c*

b. INCOME ROSTER

Line No (1)	Income type (2)	Income code (3)	This reference period (4)	Previous reference period (5) Should not have been listed	Previous reference period (5) Was last should be
1		1252	1254 1 ☐ Yes — *Mark ISS* 2 ☐ No	1256 1 ☐	
2		1258	1260 1 ☐ Yes — *Mark ISS* 2 ☐ No	1261 1 ☐	
3		1260	1262 1 ☐ Yes — *Mark ISS* 2 ☐ No	1263 1 ☐	2 ☐
4		1264	1266 1 ☐ Yes — *Mark ISS* 2 ☐ No	1267 1 ☐	2 ☐
5		1268	1270 1 ☐ Yes — *Mark ISS* 2 ☐ No	1271 1 ☐	2 ☐
6		1272	1274 1 ☐ Yes — *Mark ISS* 2 ☐ No	1275 1 ☐	2 ☐
7		1276	1278 1 ☐ Yes — *Mark ISS* 2 ☐ No	1279 1 ☐	2 ☐
8		1280	1282 1 ☐ Yes — *Mark ISS* 2 ☐ No	1283 1 ☐	2 ☐

c. [illegible] the past 4 months, that is _____, _____, _____, and _____, did . . . get income from *(Read income types in 11b column (2))*?

MARK (X) APPROPRIATE BOX IN ITEM 11b, COLUMN (4) FOR EACH INCOME TYPE LISTED.

12a. During this 4-month period, did . . . get any income from the Federal Government (that we haven't talked about)?

1284 1 ☐ Yes
2 ☐ No — *SKIP to 13a*

b. What was it called?

Anything else?

Mark (X) all that apply.

1286 1 ☐ Social Security — *Mark "1" on ISS*
1288 2 ☐ Federal Supplemental Security Income (Federal SSI) — *Mark "3" on ISS*
1290 3 ☐ A serviceman's or widow's pension from the Veterans Administration (VA) — *Mark "8" on ISS*
1292 4 ☐ Anything else — *Mark appropriate code on ISS and specify*

1294 ☐☐

13a. During this 4-month period, did . . . receive any (other) pension, disability, retirement, or survivor income (that we haven't talked about)?

1296 1 ☐ Yes
2 ☐ No — *SKIP to Check Item R8*

b. What was the source of this income?

Anything else?

Mark (X) all that apply.

1298 1 ☐ U.S. Government Railroad Retirement — *Mark "2" on ISS*
1300 2 ☐ Black Lung payments — *Mark "9" on ISS*
1302 3 ☐ Worker's Compensation — *Mark "10"* *SS*
1304 4 ☐ Payments from a sickness, accident or [illegible] insurance policy purchased on your ow[n] — *[illegible]ility Mark "13" on ISS*
1306 5 ☐ Pension from company or union — *Mark* *on ISS*
1308 6 ☐ Federal Civil Service or other Federal civilian employee pension — *Mark "31" on ISS*
1310 7 ☐ U.S. Military retirement pay (exclude paymen[ts] from the Veterans Administration) — *Mark "32" o[n]*
1312 8 ☐ National Guard or Reserve Forces retirement — *Mark "33" on ISS*
1314 9 ☐ State government pension — *Mark "34" o[n]*
1316 10 ☐ Local government pension — *Mark "35" on*
1318 11 ☐ Income from paid-up life insurance policies or annuities — *Mark "36" on ISS*
1320 12 ☐ Other or DK — *Specify and enter code from source list. If income type is not listed or DK, enter code "38" — Mark ISS.*

1322 ☐☐

CHECK ITEM R8 Is "Medicare" marked for . . . on cc item 47?

1324 1 ☐ Yes — *Mark "172" on ISS and SKIP to Check Item R23, page 8*
2 ☐ No

produced results consistent with expected differentials in net worth across subgroups. They also looked at microlevel changes in net worth using only households with fully reported wealth data and found some evidence that the dependent interview reduced the estimates of the change in net worth. More analysis of this experiment will soon be available (Weidman, King, and Williams, 1988).

The same questionnaire design issue, the dependent versus independent interview, has also occurred in the repeated measurement of industry and occupation. During the 1984 and 1985 SIPP Panels these data were collected independently during each interview even though the individual had not changed employers. This procedure acknowledges the fact that an employee's duties may change from time-to-time and allows these changes to be recorded. Sufficient change in duties can result in a change in the person's occupation classification from interview to interview even though the employer has not changed.

The independent collection of industry and occupation data has, however, several problems. Undue variation in occupation classification can result when respondent descriptions of duties vary slightly or when the interpretation of the written description varies between the clerical staff members assigning the classification codes. Research into this problem has provided some estimates of the number of times occupation and industry classifications change from interview to interview for persons with the same employer. Among individuals who reported the same employer during the first 12 months of the 1984 SIPP Panel, approximately 40 percent of these persons changed three-digit occupation codes between two consecutive interviews and 20 percent changed three-digit industry codes. In addition, only about 50 percent of persons with the same employer in all 12 months had the same occupation and only about 70 percent had the same industry code in all three waves (Kalton, McMillen, and Kasprzyk, 1986).

As a result, a modification was made in the 1986 SIPP Panel to reduce changes in occupation and industry codes resulting from random response error and clerical interpretation, and to reduce interview time. The modification introduces a "screener" question that asks if activities or duties have changed during the past 8 months. A negative response eliminates the detailed occupation and industry questions. The occupation and industry classifications are then brought forward from the previous interview.

It is important to note that while this change was made for the 1986 SIPP Panel, industry and occupation data from the 1985 SIPP Panel, collected during the same time period, were still collected independently each wave, giving rise to a natural experiment embedded in the two panels. These data have not yet been analyzed.

3.3 Missing Interview Questionnaire

In panel surveys respondents may miss one or more interviews. When this occurs it is possible that collecting retrospective data for missed interviews may alleviate the problem of nonresponse. Other errors, however, such as recall error may be introduced into the survey. In order to determine the feasibility of obtaining retrospective information covering periods of missed interviews, a new section was added to the questionnaire, a section called the "missing wave." This section of the questionnaire was used to gather information at

wave (i+1) for interviews missing in wave (i) conditioned on the fact that data were available from wave (i-1). The missing wave section of the questionnaire did not contain all missing questions but rather a very reduced set of questions concerning labor force status, receipt of income from assets, jobs, and program participation.

In principle, this appears to be a reasonable approach compared to imputation when handling a specific type of missing data problem in a panel survey. Huggins (1987a) evaluated the use of this sequence of questions and concluded that the small number of transitions observed for specific income types did not justify the respondent burden and cost of asking the additional questions, since comparable methods, such as a direct substitution imputation, were available.

3.4 Annual Roundup

The SIPP obtains monthly data for a 4-month reference period from a variety of income sources. The relatively short reference period and repeated interviews every 4-months should result in better estimates of income received during a calendar year. One topical module on the SIPP, however, concentrates on direct questions on annual amounts received (using the W-2 form obtained from the employer)--the annual round-up/tax topical module. These questions have two purposes: 1) to provide alternative estimates of annual calendar-year income for a selected group of income sources, and 2) to provide information to guide imputation models for item nonresponse for individuals not reporting in one or more interviews. The first issue which needs to be addressed is how estimates obtained by summing monthly amounts collected in the core data compare with the direct question on annual earnings. Preliminary findings are reported in a Census Bureau memorandum (U.S. Bureau of the Census, 1988a). One rather discouraging result was that persons who had imputations in the monthly core questions were also very likely to be nonrespondents on the annual sequence of questions. Much more work on this topic is necessary before models of annual earnings can be delivered.

3.5 The Collection of Employer-Provided Benefits

In recent years interest in employer contributions to health insurance, retirement, and life insurance plans have become an important focus of national attention. Since one of the goals of the SIPP is to provide improved measures of economic well-being, research was initiated on the collection procedures and questionnaire design appropriate for obtaining data of this type. A small study was conducted with the last rotation group of the last interview of the 1985 Panel (August 1987). The aim of the study was to determine the feasibility of obtaining the amount of the employer and employees contributions to health insurance, pension, and life insurance plans. One-half of the sample cases in the last rotation group were used in this study. A short questionnaire on these topics was sent to the employers of individuals in the survey upon authorization from the individual respondent. The two principal issues surrounding this study are 1) would respondents sign a form authorizing the Census Bureau to contact their employers, and 2) would the employers send the information to the Census Bureau with the approval of their employees. The use of a signed released procedure had been implemented in other surveys in particular, the National Medical Care Utilization and Expenditure Survey (NMCUES) and the National Medical Care Expenditure Survey (NMCES). This small study, however, was the SIPP's first attempt at such methodology. Obviously, the

analysis will center around respondent cooperation in signing releases, employer response rates, missing data rates, and cost. Carmody, Fischer, and Meier (1988) provide a description of the study and some preliminary analysis.

4. Data Collection

Four topics affecting data collection in the SIPP are discussed below: 1) respondent rules; 2) data collection mode; 3) length of reference period; and 4) rules for following movers.

4.1 Respondent Rules

When interviewing households with more than one member, a problem which must be addressed is the extent to which proxy responses are acceptable. Since not everyone may be present at the time of the interview, both time and money can be saved by asking another household member about persons who are not present. The difficulty with this is that along some dimensions of the survey instrument, the proxy report may result in less accurate data than the self-report. Kalton, Kasprzyk, and McMillen (1988) provide a discussion of this issue in the context of panel surveys.

A formal test of respondent rules, conducted in the ISDP, compared the quality of reporting in a treatment group where proxy interviews are accepted from any household member who felt qualified to answer for a missing person with a treatment group where proxy interviews are not permitted except for extreme situations (respondent physically or mentally incapable, unable to speak English, away from the household during the entire interviewing period, etc). About 85 percent of adults interviewed in the self-response rule households were self-respondents and about 65 percent were self-respondents in the usual or proxy-response rule households. Thus, the implementation of the self-response rule resulted in approximately 20 percent more self-interviews than the other treatment (Coder, 1980).

Refusal rates were slightly higher for the self-response treatment and the percent of households interviewed was slightly higher for the proxy-response treatment. The differences, however, were too small to give insight into which rule should be preferred. Person noninterview rates in households where at least one other adult was interviewed were higher under self-response rules than under usual response rules. Differences between treatment groups in reported income recipiency rates also appeared to be small and unaffected by the response rule, and combined "don't know" and "refusal" rates for income amounts of various income types were not consistently lower under the self-response mode.

Under the self-response rules, records were used more often by persons when answering wages and salary questions, and response rates for hourly wage rates were higher; but in general the evidence for either set of response rules was not conclusive. Thus, as a result of these findings, estimated costs for using a self-response rule (4-to-6 percent higher than the proxy rule), and the implementation of a "call back" procedure to obtain certain critical information unavailable at the time of the interview, the SIPP respondent rules now allow proxy interviews to be taken.

The respondent rules adopted for the SIPP are that adults present at the time of the interview report for themselves while proxy informants are accepted for absent adults. A hierarchy of proxy informants has been established for the SIPP so that a spouse is always the first choice as a proxy; the second-level proxy is the adult who was the proxy at the previous interview; the third-level proxy is an individual who was proxy at any other interview; and finally first-time proxy is accepted.

Observation of self-proxy rates on a cross-sectional basis over the course of the panel reveals little variation--63 percent to 67 percent of the respondents at each interview report for themselves. However, Kasprzyk and McMillen (1987) report a somewhat different picture when considering self-proxy reporting patterns over the length of the panel. They found that only 40 percent of the individuals who participated in all eight interviews of the panel were self-reporters at each interview. Another 19 percent of the individuals had only 1 or 2 proxy interviews conducted, about 11 percent never reported for themselves. Except for a specific problem related to the measurement of state-to-state transitions (Weidman, 1986) and one of labor earnings for prime-aged males (Hill, 1987a), no significant data analysis addressing the self-proxy reporting issue has taken place. In view of the extent of proxy reporting in the SIPP, the nature and quality of self-proxy responses during the panel should be addressed sometime in the near future.

A related problem is the response rule for college students. Students are usually considered members of their parents' households until they establish a permanent residence elsewhere. Thus, the usual procedure for students living away from home while attending school is to treat them as household members who are temporarily absent and obtain proxy interviews from other members of their parents' household. In order to measure the accuracy of information taken from proxy interviews for students living away from home, one interview during an ISDP field test was first obtained by proxy at the parents' household and then by self-interview at the student's school residence. The results of this study are described by Roman and O'Brien (1984). The analysis presented is limited due to flaws in the administration and implementation of the test. The authors observed, however, that quite often a proxy cannot identify a particular source of student income and, even if they can identify it, they are more likely to respond "don't know" to the particulars about that source. They also noted that the larger the income or expense, the better the proxy response becomes.

4.2 Data Collection Mode

The SIPP has conducted most interviews (approximately 95 percent) face-to-face. Because of the rising costs of face-to-face interviews, the Census Bureau is considering the possibility of conducting a substantially larger number of SIPP interviews by telephone. Considerable disagreement existed among the staff working on the SIPP over the practicality of using the current questionnaire with a telephone interview, since the questionnaire is long, complex, and relies on numerous "check items" which route individuals through the form depending on a variety of statuses. Some also felt that the sensitive nature of the topics covered, income and related matters, would result in large amounts of missing data. In order to understand both the interviewers and respondents reaction to the telephone collection mode in the context of SIPP, a telephone interview pretest was conducted in June 1985. The pretest was conducted in 2 of the Census Bureau's Regional Offices with a sample of 280 households.

Refusal rates (about 2.5 percent) and noncontact rates (about 11 percent) were within staff's expectations. No unexpectedly high nonresponse rates were observed (U.S. Bureau of the Census, 1986a).

Following this, a SIPP National Telephone Test took place from August to November 1986 and February to April 1987; the purpose of the test was to study the large-scale use of warm telephoning in SIPP and to learn whether people are willing to furnish data by telephone for two interviews in a row. Households within 50 percent of the segments were designated as maximum telephone interview cases; the remaining 50 percent were maximum personal visit cases. Interviewers conducted almost all of the telephone interviews from their homes. Gbur and Durant (1987) and Carmody, Fischer, and Meier (1988) report preliminary results from the first phase of the experiment.

They indicate that household response rates did not seem to be seriously affected by the use of the telephone and person nonresponse rates were comparable by mode. Item nonresponse rates were only slightly affected by telephone interviewing; additional analysis is forthcoming.

4.3 Length of Reference Period

The ISDP focussed on data collection techniques designed to improve the reporting of cash and noncash income, and as such the length of the reference period for most survey items was an important design decision.

This issue was addressed twice during the ISDP. First a single interview using a 6-month recall period was compared with two consecutive interviews, both using 3-month reference periods. Second, an experiment was conducted comparing reported property income amounts using a 3-month recall versus those with a 6-month recall period. Thus, the receipt of property income was collected from the full sample, but the sample was randomly split, with one-half the sample reporting amounts of property income received for a 3-month period, while the remaining half of the sample reported amounts obtained from property income for a 6-month interval.

Olson (1980) describes some analyses conducted on the first experiment. Not surprisingly, the proportion of respondents reporting some positive amount of income in the initial 3-month reference period is higher for the 3-month reference period group than for the 6-month reference period group; that is, using a 6-month recall period understates the proportion of income reported in earlier periods. This pattern held for a number of specific sources of income such as wages, Aid to Families with Dependent Children, and unemployment compensation. These findings, though not definitive, support the presumption that longer recall periods increase chances of omission due to memory loss. Other analysis showed that the number of sources of income reported per household in the first 3 months of the 6-month reference period was lower than for the corresponding time using a 3-month reference period. Analyses of the second experiment were not conducted due to the withdrawal of funding for the development program.

The results of the first experiment along with additional ISDP experience led to a 4-month recall period for the SIPP; this decision maintains cost at the appropriate budget level while trying to maintain satisfactory data quality. The problem of recall error has not been studied systematically in the SIPP.

The SIPP Record Check Study (section 6) may provide insight into recall problems in the reporting of monthly income from selected income sources. The Statistical Methods Staff (U.S. Bureau of the Census, 1986b) reviewed September 1983 data to determine whether the number of months between the occurrence of the event and the reporting of the event affects the reported values. The analysis found a "recall lag" effect. This result is not definitive, however, since there were few changes/transitions within the interview wave.

4.4 Rules for Following Movers

An important design feature in the ISDP and now the SIPP is that all persons in a sample household at the time of the first interview remain in sample during the 2 1/2-year period of the panel; this rule holds even if one or more persons should move to a new address. For cost and operational reasons, face-to-face interviews are conducted at new addresses that satisfy some geographic constraint--in the ISDP, the address had to lie within 50 miles of an ISDP primary sampling area; while in SIPP, the address must lie within 100 miles of a SIPP primary sampling area.

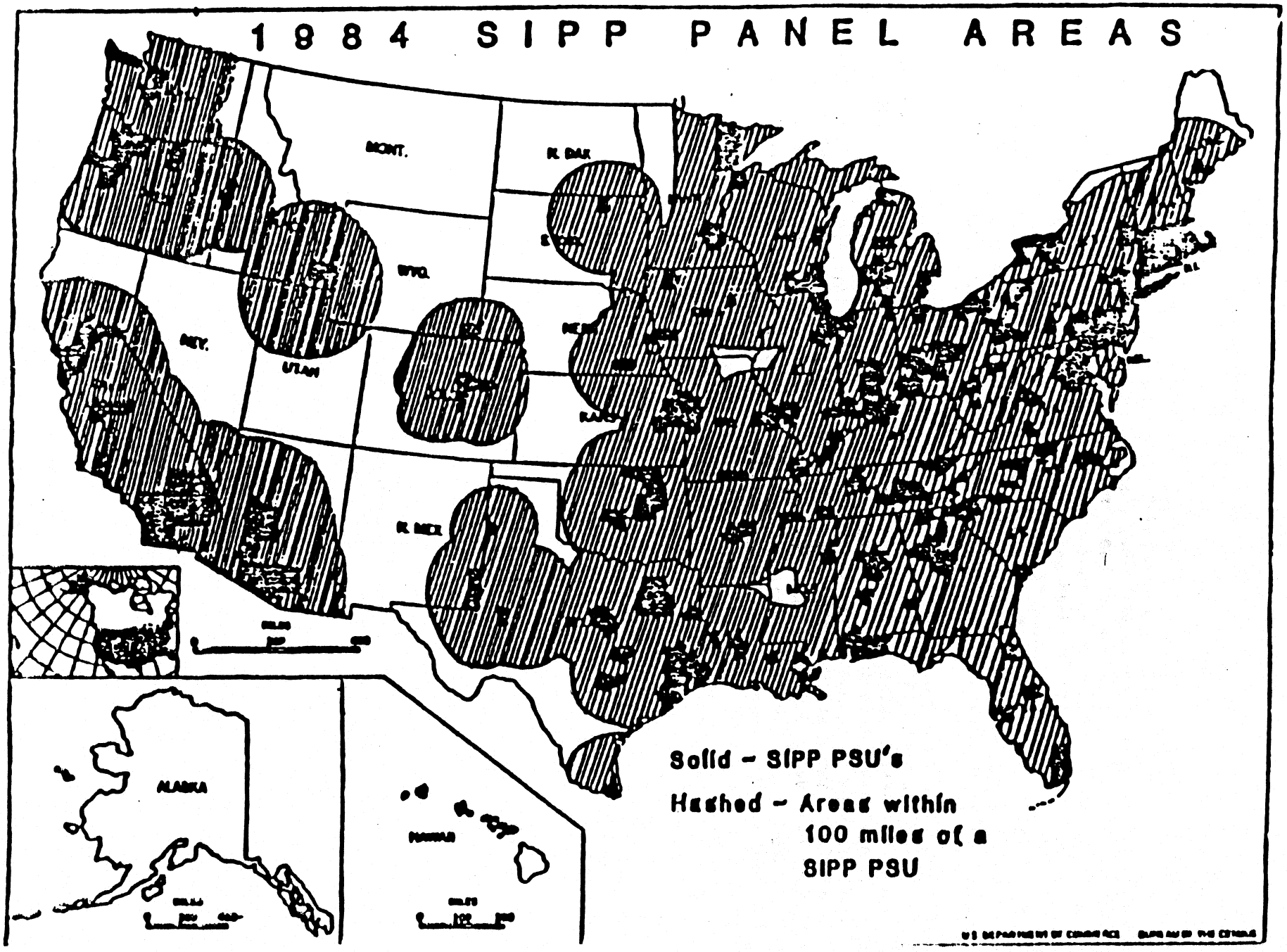
For each panel a sample of addresses is selected and individuals are identified at these addresses at the time of the first interview. After the first interview, the sample is no longer address-based but rather person-based, consisting of all individuals enumerated during the first interview. Thus, these people and anyone with whom they share living quarters ("new entrants") are interviewed in subsequent interviews.

Figure 5 shows the SIPP sampling areas for the 1984 Panel. The hash-marks illustrate the "within 100 miles of a SIPP primary sampling area" rule. Approximately 96.5 percent of the U.S. population lies within the area of the SIPP following rules. As a consequence, the rule does not appear to be very restrictive. When SIPP sample individuals move outside of the hash-marked areas, interviewers are instructed to conduct telephone interviews when possible.

During the ISDP two issues concerning movers were important: (1) the production of cross-sectional point-in-time estimates at each interview; and (2) the costs associated with following movers. Huang (1984) presents several unbiased base weights for cross-sectional estimates of the noninstitutionalized population when the sample contains movers. He associates observations at any given point in time with the known inclusion probabilities of the original sample households. Two approaches are described: 1) a multiplicity approach, which depends on the number of ways that a new household can be included in the sample; and 2) a "fair share" approach which assumes all household members contribute equally to their household. The SIPP as well as the ISDP adopted the "fair share" approach.

The issue of costs was addressed by a "Mover's Cost Study." This study was to shed some light on the data collection costs resulting from following persons to their new addresses. White and Huang (1982) describe the study and provide some results based on the movers procedures adopted for the field test. They found that the number of eligible households for interview increased 1.8 percent as a result of following movers during a one-year time period. They also found that movers represented about 22 percent of the total sample after 15 months, and that during this period of time the number of interviewing hours

Figure 5


1984 SIPP PANEL AREAS
MONT.
N. DAK.
WYO.
NEV.
UTAH
N. MEX.
ALASKA
Solid - SIPP PSU's
Hashed - Areas within
100 miles of a
SIPP PSU

increased by 7 percent and the number of miles charged by interviewers increased by 11.4 percent.

Jean and McArthur (1984) discuss data collection issues in the SIPP as they pertain to movers and offer recommendations to improve coverage in future SIPP panels. Kalton and Lepkowski (1985) also discuss the procedures for following movers adopted in SIPP, and propose a research program aimed at measuring the extent of noncoverage from various sources and its concentration in particula subgroups. More recently, Jean and McArthur (1987), considering five waves of SIPP data, report that, among persons who moved sometime after the first interview (that is, between Waves 2 and 5), 69 percent completed all five interviews, 23 percent did not complete the fifth interview, and 9 percent were interviewed in the fifth wave but were missing at least one intervening interview.

At this time there are no plans to review the operational decisions of whom to follow and where to follow. The rules for following movers can be assessed from the coverage point of view, and even though minor modifications to the rules are possible in order to improve coverage (such as following children under the age of 15 who are no longer living with the original sample person), there is no immediate intention to do so. Rather, the issue of concern in the future revolves around the interviewer's ability to find a mover and conduct the interview. This is essentially a question of assessing whether the success rate in finding movers can be improved and whether nonresponse adjustment factors can be developed which compensate for the mover population who can not be traced or who refuse to participate in the survey.

5. Concepts, Design, and Estimation

During the ISDP and continuing with the SIPP program, significant research activity has taken place in the area of conceptualizing annual units of analysis using subannual data and the statistical estimation of these concepts. The treatment of nonresponse in panel surveys has also been a topic of study; research interest has been evident in the three areas of nonresponse in a panel study--unit, wave, and item. Finally, estimation techniques to reduce sampling error and methods to sample subgroups have also been under study in the ISDP and SIPP programs.

5.1 Longitudinal Concepts

Annual family and household statistics are important indicators of the Nation's economic well-being. The SIPP collects monthly data reflecting changes in the composition of households; these data allow the development of a l household statistics which reflect actual household composition experienc ring the year, unlike current household statistics which ignore intra-year anges n household composition. The construction of annual units of analys whet r they are households, families, or program units, raises methodologi iss concerning longitudinal weights and imputation techniques. The main sue however, conceptual. Given intra-year composition change, how shoul nnua measures reflect change in household composition? That is to say, h shou households and families be defined which account for survey measurer s at o or more points in time?

Analysts at the Census Bureau have given considerable thought to the question of defining households and families over time (McMillen and Herriot, 1985; Citro, 1985). Empirical research to examine several definitions of longitudinal households and measures of annual income status and family type has been reported by Citro, Hernandez, and Herriot (1986) and Citro, Hernandez and Moorman (1986). The empirical research emphasized four concepts: 1) a household is the same over time if it has the same reference person; 2) a household is the same over time if it has the same principal person (this definition differs from the first in its treatment of married-couple households for which the reference person may be either the husband or wife, but the principal person is always the wife); 3) a household is the same over time if it has the same reference person and is the same family type over time; and 4) a household continues over time if it has the same reference person, is the same family type, and has the same membership size.

This research has provided preliminary indications that the choice of definition does not appreciably affect annual measures of low income status or of households by type. If this finding does not change after additional research, considerations, such as ease of implementation and operational simplicity, will be the determining factors in choosing a longitudinal household definition.

5.2 Statistical Estimation for Longitudinal Concepts

Research on estimation for longitudinal concepts has proceeded along two paths--longitudinal person estimation and longitudinal household (family or program unit) estimation. The work on person estimation includes the calculation of selection probabilities to yield unbiased longitudinal estimates of individual characteristics and the use of controls in additional stages of estimation (Judkins et al., 1984). A refinement of this work and a description of the method proposed to produce longitudinal weights for person analysis covering the first three SIPP interviews has been reported by Kobilarcik and Singh (1986).

Kobilarcik and Singh define the longitudinal universe as the noninstitutional population (excluding military barracks) on December 1, 1983, the midpoint of the Wave 1 interview months. The sample from the longitudinal universe consists of eligible persons living in the selected living quarters at the time of the first interview. "Interviewed" persons for purposes of this estimation procedure are 1) those who responded to each of the first three SIPP interviews, and who during the first interview lived in a household in which all eligible members responded to the interview; and 2) those individuals who resided in a Wave 1 interviewed household, but during the second or third interview died or moved to an ineligible address.

Thus, noninterviewed persons in the estimation procedure are those who at the time of the first interview lived in a household in which at least one household member failed to respond to the first interview, and those who resided in a Wave 1 interviewed household but failed to respond at the second and/or third interview. All persons classified as interviewed are assigned positive weights. Weights for this universe are derived in the usual way, using the reciprocal of the probability of selection, calculating an adjustment for noninterviews, and adjusting to demographic population controls. The nonresponse adjustment has two phases, an adjustment first for household nonresponse and then for person nonresponse, the latter using information collected during the first interview.

A conceptually similar approach has been adopted for the panel (eight interviews) microdata file. The longitudinal universe for the panel consists of eligible persons living in the selected living quarters at the time of the first interview. A "panel" or "cohort" weight for the eight-interview file is developed by treating only those individuals who responded at all eight interviews as "interviewed" persons. The cohort weight serves to represent the United States noninstitutional population (excluding military barracks) as of November 1983. This weight is useful for analyzing data over the 2 1/2-year time period with the analytic unit being the individual.

A similar point of view has been adopted for the estimation of calendar-year characteristics for the individual. SIPP, as a panel, obtains 2 1/2 years of monthly income data. As such the development of 2 successive years of calendar-year income estimates is possible. Thus, in addition to the "panel" weight, two other weights exist on the panel file--the weights applicable to the development of calendar-year income estimates from each of the 2 calendar years, 1984 and 1985.

Specific details can be found in a memorandum from Statistical Methods staff (U.S. Bureau of the Census, 1988b). Generally, the idea is to use all individuals (initial sample people or new entrants) who have provided data for all interviews during the time period of interest. Once again, the two weights are cohort weights for 1984 and 1985, and represent the civilian noninstitutionalized population on January 1, 1984, and January 1, 1985. Thus, individuals with positive weights are those who participated in the survey throughout 1984 or throughout 1985.

The panel file of eight interviews contains a data record for everyone who has ever been in the survey. Thus, the availability of three weights although relatively easy to understand, may result in some confusion for analysts of the SIPP panel data, since the set of positively weighted individuals will change depending on the time period the analysis is to cover.

The topic of longitudinal household (family or program unit) estimation has also been studied. Several approaches to this issue were reported by Ernst, Hubble, and Judkins (1984) and more recently by Ernst (1988). The latter work describes why weighting by the reciprocal of the probability of selection does not, in general, work for longitudinal household and family estimates, and presents a class of weighting procedures which can accomplish this task. Ernst, furthermore, describes the difficulties that can arise in applying these weighting procedures because the information necessary to create the weight may not be available. Ernst also presents conditions which, if satisfied, by the longitudinal concept, are sufficient for there to exist a weighting procedure that avoids these problems. Finally, he discusses procedures for adjusting longitudinal concepts for nonresponse and for controlling demographic variables to independent estimates. These procedures may need to be treated differently for estimation of longitudinal concepts than for cross-sectional concepts.

The topics discussed in this section have been under development for an extended period of time. Longitudinal household concepts for SIPP were first discussed by Griffith (1978) and Lane (1978). Some elementary thoughts about the estimation of the concepts were discussed by Kasprzyk and Kalton (1983). Empirical research comparing several concepts was not undertaken until 1984/1985 with Citro's work as an American Statistical Association/National Science Foundation/

Census Bureau Research Fellow. Finally, with the end of the development of the first panel file and its release to the public, we expect further evaluations of the estimation procedures and of the usefulness of longitudinal household concepts. These evaluations will likely consider the quality of the estimates, the reasonableness of the results from a "real" world point of view, and the effect of attrition on estimates of the number and characteristics of longitudinal households. After ten years of discussing whether and how to analyze longitudinal households, we are now in a position to evaluate empirically concepts and estimation procedures. Work on these topics will occur during 1988-1989.

5.3 Nonresponse and Imputation

Nonresponse in longitudinal surveys can be treated from either the cross-sectional or longitudinal point of view. Either treatment is valid depending on the uses of the data. If a longitudinal analysis is conducted, then treating nonresponse from the longitudinal perspective is more desirable since it reflects the survey design. This point of view, because of the repeated interviews, often provides information which is highly correlated with the missing data--the same information measured at different points in time as well as information on patterns of behavior and transitions from one state to another. Thus, under this perspective, nonresponse is not viewed as nonresponse in a set of unrelated observations but as nonresponse in a set of variables with a logical dependency between two or more points in time. For example, in the SIPP, income data obtained at time t-1 or time t+1 can be used to impute for missing income data at time t. This view adds considerable information to the data set for the treatment of nonresponse and justifies matching waves as quickly as possible to treat nonresponse from a longitudinal perspective.

If nonresponse in a longitudinal survey is treated from a cross-sectional perspective, each wave is treated as a separate survey. This has practical advantages in that the release of wave data may occur more quickly than if the separate waves were first linked, and linkage and editing problems resolved. A disadvantage is that records with imputed data will be inconsistent from wave to wave because data processing and estimation procedures are implemented independently from one time to the next. Despite the inconsistencies at the micro-record level, changes in aggregates from one wave to another can be investigated.

An additional complication to the treatment of nonresponse comes from the fact that, in SIPP, unit nonresponse can be measured in several ways (Chapman, Bailey, and Kasprzyk, 1986). The typical way is to consider the total number of eligible households assigned including the "Type A" noninterviews (household noninterviews including refusals, no one at home, etc.) for Wave 1 as denominator. The numerator is then the total number of Type A's in the survey.

In SIPP, an additional form of unit noninterview exists because survey procedures call for following all people who lived at the sample address at the time of the first interview. Thus, a "Type D" noninterview household is defined as a household of one or more original sample persons who cannot be followed to their new address(es) or moved beyond 100 miles of a SIPP PSU. Table 1 provides cumulative Type A and Type D household nonresponse rates by wave for the 1984 to 1987 SIPP Panels. As mentioned earlier, by nature of its design, the SIPP should expect its cumulative nonresponse rate to increase after each

TABLE 1
HOUSEHOLD NONINTERVIEW RATES AND SAMPLE LOSS 1/

1984 SIPP Household Noninterview Rates and Sample Loss

Wave	Type A Rate	Type D Rate	Sample Loss 2/
1	4.9%	--	4.9%
2	8.3%	1.0%	9.4%
3	10.2%	1.9%	12.3%
4	12.1%	2.9%	15.4%
5	13.4%	3.5%	17.4%
6	14.9%	4.1%	19.4%
7	15.6%	4.9%	21.0%
8	15.8%	5.7%	22.0%
9	15.8%	5.7%	22.3%

1985 SIPP Household Noninterview Rates and Sample Loss

Wave	Type A Rate	Type D Rate	Sample Loss 2/
1	6.7%	--	6.7%
2	8.5%	2.1%	10.8%
3	10.2%	2.7%	13.2%
4	12.4%	3.4%	16.3%
5	14.0%	4.1%	18.8%
6	14.2%	4.8%	19.7%
7	14.4%	5.2%	20.5%
8	14.4%	5.5%	20.8%

1986 SIPP Household Noninterview Rates and Sample Loss

Wave	Type A Rate	Type D Rate	Sample Loss 2/
1	7.3%	--	7.3%
2	10.8%	1.5%	13.4%
3	12.6%	2.3%	15.2%
4	13.8%	3.0%	17.1%
5	15.2%	3.7%	19.3%
6	15.2%	4.3%	20.0%
7	15.3%	4.8%	20.7%

1987 SIPP Household Noninterview Rates and Sample Loss

Wave	Type A Rate	Type D Rate	Sample Loss 2/
1	6.7%	--	6.7%
2	11.1%	1.5%	12.6%
3	11.5%	2.6%	14.2%
4	12.3%	3.3%	15.9%

1/Type A noninterviews consist of households occupied by persons eligible for interview and for whom a questionnaire would have been filled if an interview had been obtained. Reasons for Type A noninterview include: no one at home in spite of repeated visits, temporarily absent during the entire interview period, refusal, and unable to locate a sample unit.

Type D noninterviews consist of households of original sample persons who are living at an unknown new address or at an address located more than 100 miles from a SIPP PSU, provided a telephone interview is not conducted.

2/The sample loss rate consists of cumulative noninterview rates adjusted for unobserved growth in the Type A noninterview unit (created by splits).

interview--and it does--but the wave-to-wave change in sample loss decreases during the course of the panel. See Nelson, Bowie, and Walker (1987) for a more complete review of unit nonresponse in the SIPP, including comparisons of the SIPP sample loss with those observed in other panel surveys and a discussion of the methods Census Bureau staff are using to maintain respondent cooperation. With nonresponse accumulating during the panel, some concern about the effectiveness of the household nonresponse compensation procedures exists. In particular, the issue is the selection of the most effective weighting classes for nonresponse adjustment. Petroni and King (1988) describe a study which weights the sample in two ways in order to see the effectiveness of the cells chosen as nonresponse adjustment cells. Their work, even though preliminary, suggests that it might be helpful to include monthly household income, metropolitan/nonmetropolitan, and a further breakdown of the race and Spanish-origin cells.

Another way of viewing response rates in the SIPP is to look at them on a "person" basis as opposed to a household basis; that is, consider sample loss in terms of the reduction in the numbers of initially interviewed sample persons over the time these individuals were eligible for interview. Table 2 and table 3, taken from Kasprzyk and McMillen (1987), provide a summary of response patterns observed during eight SIPP interviews. A detailed accounting of response patterns is available in an internal Census Bureau memorandum (U.S. Bureau of the Census, 1988c). This memorandum also presents tables giving the distribution of the reasons for noninterviews and characteristics of persons who leave the 1984 Panel sample.

Under either way of measuring nonresponse--rates based on household or rates based on persons--it is clear that, in a longitudinal survey, a missing data problem exists which is different from the cross-sectional missing data problem. This is the problem of wave nonresponse--individuals or households who miss one or more (but not all) interviews--and its treatment.

The amount of missing data for an individual with wave nonresponse is typically greater than that encountered for records with item nonresponse. Data available from completed waves of interviewing, however, provide more information about the nonresponding unit than is available for total nonrespondents. Thus, nonresponse compensation strategies may include weighting, imputation, or a combination of both. Kalton, Lepkowski, and Lin (1985) discuss this issue and empirical findings in the context of the ISDP. This work made it clear that the choice between weighting and imputation for missing data of this type is far from obvious. Kalton (1986) and Kalton and Miller (1986) further refine the understanding of this problem and conclude that imputation can distort some forms of estimates and that weighting may be the preferred solution for large subclasses when the reduction in effective sample size is tolerable. They caution, however, that imputation may be better for estimates based on small subclasses when the loss of sample is important. In the case of a three-interview longitudinal SIPP file the difference in sample size between weighting and imputation is not substantial, and, consequently the weighting approach is the safer general-purpose solution; however, in an eight-interview longitudinal SIPP file the choice is by no means obvious. Finally, Lepkowski (1988) after further empirical research concludes that a specific strategy for wave nonresponse can only be developed after consideration of such factors as the major survey design objectives, the panel design, and the distribution of wave nonresponse patterns. He provides criteria to be considered in developing

TABLE 2
Response Patterns for Original Sample Persons
(100-level) 1/

Response every interview (8 interviews)		Number	Percent
Pattern:	XXXXXXXX	32192	73.0%
Attrition Cases		8173	18.5%
Patterns:	XXXXXXX0	623	1.4%
	XXXXXX00	802	1.8%
	XXXXX000	919	2.1%
	XXXX0000	1149	2.6%
	XXX00000	1259	2.9%
	XX000000	1603	3.6%
	X0000000	1818	4.1%
All other Patterns		3723	8.5%
Total		44088	

1/The universe for the table consists of all persons eligible for eight interviews in the 1984 SIPP Panel and for whom a personal interview was conducted (either self-or proxy-interview) during the first wave of the 1984 SIPP Panel. The symbol "X" represents a successful interview and the symbol "0" represents no interview (either no household interview or no personal interview).

TABLE 3
Number of Missing Interviews for Original Sample Persons
(100-level) 2/

Number of Interviews Missing	Number of Persons	Percent
0	32192	73.0%
1	2794	6.3%
2	1377	3.1%
3	1283	2.9%
4	1396	3.2%
5	1482	3.4%
6	1746	4.0%
7	1818	4.1%
Total	44088	100.0%

2/The universe for the table consists of all persons eligible for eight interviews in the 1984 SIPP Panel and for whom a personal interview was conducted (either self-or proxy-interview) during the first wave of the 1984 SIPP Panel.

missing data strategies and concludes that weighting strategies appear to be preferable for compensating for wave nonresponse.

Item nonresponse, as with unit nonresponse, can be viewed from the cross-sectional or longitudinal dimension. Item nonresponse typically refers to missing data items in an otherwise completed interview. It provides a good illustration of the fact that there is nothing theoretically special about longitudinal imputation for item nonresponse. Longitudinal imputation for item nonresponse is simply imputation for item nonresponse using auxiliary data from a larger data base, using longitudinal data elements as well as cross-sectional ones.

Discussions of the levels of item nonresponse in the SIPP have occurred periodically at the meetings of the American Statistical Association (Coder and Feldman, 1984; Lamas and McNeil, 1984; McMillen and Kasprzyk, 1985). These reports have focused on cross-sectional item nonresponse rates. One general observation common to these papers is that for "core" data from the SIPP, the levels of item nonresponse are low. In addition to the papers cited above, levels of item nonresponse can be found in the U.S. Bureau of the Census, Current Population Reports, Series P-70, "Economic Characteristics of Households in the United States." Table 4 provides a summary of SIPP item nonresponse rates for each calendar quarter of 1984 compared to the March 1985 Current Population Survey.

The concept of cross-sectional item nonresponse based on data obtained in one interview can be extended to a longitudinal concept that combines the nonresponse experience for successive interviews. This has been done for the first three observations in the 1984 SIPP; the results for a selected group of income types are shown in table 5. The rates in this table are based on the total number of persons reporting receipt of the specified income type at any time during the 12-month period. The first column shows the percent of all income recipients that reported amounts for all months during which the income source was received. The other columns indicate situations in which amounts were not reported in one or more, one or more but not all, and all months of recipiency. The right-most column showing the proportion of cases for which no income amount was reported indicates that only in a small number of cases was no information available.

The treatment of item nonresponse in the longitudinal context was described by Heeringa and Lepkowski (1986) and Kalton and Lepkowski (1983). Heeringa and Lepkowski empirically compare a simple longitudinal imputation method, longitudinal direct substitution (a value of a nonmissing item is substituted from one time period to another when the same item is missing), with a cross-sectional hot-deck scheme. Not surprisingly, they demonstrated that the direct substitution method for longitudinal imputation understates change. They concluded, however, that this may be preferable to the gross overstatement of change resulting from the use of the cross-sectional hot-deck method.

Other imputation work focussing on model development has been conducted by Huggins and Weidman (1986a, 1986b). Models which impute missing reponse patterns have also been investigated (Samuhel and Huggins, 1984: Huggins, Samuhel, and Weidman, 1985).

Table 4. Item Nonresponse Rates for SIPP and March 1985 CPS, for Selected Income Types

Income Type	SIPP 1984 1st Quarter Monthly Average	SIPP 1984 2nd Quarter Monthly Average	SIPP 1984 3rd Quarter Monthly Average	SIPP 1984 4th Quarter Monthly Average	March 1985 CPS
Wage and Salary	7.2	7.5	7.5	7.6	18.9
Self-Employment Income	16.8	16.2	16.0	16.1	26.5
Social Security Income	7.6	8.4	8.1	8.4	19.9
Federal Supplement Security Income	10.8	11.6	11.7	12.3	21.3
Aid to Families with Dependent Children	6.1	6.9	6.5	5.5	16.0
Unemployment Compensation	10.1	13.6	10.4	12.7	21.8
Company or Union Pension	13.9	14.0	12.8	14.7	24.0
Food Stamp Allotment	5.2	6.3	6.7	6.6	13.7
Veterans' Compensation or Pension	11.3	11.2	11.9	13.5	18.3

1. U.S. [illegible] Census, Current Population Reports, Series P-70, No. 3, Economics Characteristics of Households ... [illegible] United States: First Quarter 1984, U.S. Government Printing Office, Washington, D.C. 1985

2. U.S. Bureau of the Census, Current Population Reports, Series P-70, Economics Characteristics in the United States: Second Quarter 1984, U.S. Government Printing Office, Washington, D.C. 198[illegible].

3. U.S. Bureau of the Census, Current Population Reports, Series P-70, Economics Characteristics of Households in the United States: Third Quarter 1984, U.S. Government Printing Office, Washington, D.C. 1985

4. U.S. Bureau of the Census, Current Population Reports, Series P-70, Economics Characteristics of Households in the United States: Fourth Quarter 1984, U.S. Government [illegible]ing Office, Washington, D.C. 1986

TABLE 5 -- Longitudinal Item Nonresponse Rates for Amounts of Selected Income Types: 1984 SIPP Panel 12-Month Summary 1/

(percent)

INCOME TYPE	ALL AMOUNTS REPORTED	ONE OR MORE AMOUNTS NOT REPORTED	ONE OR MORE BUT NOT ALL AMOUNTS NOT REPORTED	NO AMOUNTS REPORTED
Hourly Wage Rate	83.0	17.0	9.0	8.0
Social Security.	82.8	17.2	13.1	4.1
Private Pension.	78.8	21.8	13.6	8.2
AFDC	91.0	9.0	5.6	3.4
Food Stamps.	91.9	8.1	6.2	1.9
Unemployment Compensation.	87.9	12.1	4.0	8.0
Federal SSI.	88.0	12.0	7.6	4.4

NOTE: These rates are based on the total number of persons with recipiency in one or more of the 12-months. Also these rates do not reflect imputations made to type Z person noninterviews.

1/This table, prepared by John Coder of the Census Bureau, is taken from SIPP Working Paper Series No. 8601, "Some Aspects of the Survey of Income and Program Participation," compiled by Daniel Kasprzyk and Roger A. Herriot.

5.4 Sampling Error Reduction through Estimation Techniques

Two methods for reducing sampling error through estimation techniques are under study: composite estimation and the use of administrative records in SIPP estimation.

Composite estimation is a technique that combines estimates from the current and previous time periods with the goal of improving the precision of cross-sectional estimates by taking advantage of the correlations between responses for the same analytic units at different time periods. Composite estimation is particularly effective when the correlations are high, which is likely to be the case for many important data items in the SIPP. Chakrabarty (1986) has conducted a preliminary review of the types of composite estimates appropriate for the SIPP data structure. The content of the survey has not been sufficiently stable during the first few years of the SIPP to seriously consider adoption of a composite estimator.

Another approach to variance reduction is through the use of administrative records for post-stratification. Currently, cross-sectional estimation procedures for SIPP make use of a second-stage adjustment to increase the precision of estimates by ratio adjusting collection month and reference month estimates to population estimates. However, the Census Bureau has access to some Internal Revenue Service (IRS) and Social Security Administration (SSA) files which can be used to produce detailed age, race, and sex distributions by adjusted gross income. The issue, which we have just begun to explore, is how these administrative data can be used for post-stratification to improve estimates of mean and median personal and household income as well as the estimates of the deciles of the personal and household income distributions.

The first phase of this research (U.S. Bureau of the Census, 1987) will estimate the reductions in variances of SIPP estimates by using the IRS data as auxiliary variables in the estimation procedure. The procedure being studied has been advocated by Herriot (1983) and Scheuren (1983). In the SIPP study the estimation method will involve a ratio adjustment of SIPP estimates at the second stage of estimation in cells defined by age + race + sex + "income" where "income" is adjusted gross income as reported to the Internal Revenue Service.

Controls are prepared from a 1-percent sample of the 1984 IRS file matched with age, race, and sex characteristics from the SSA Summary Earnings Record (a file containing individuals lifetime covered earnings, up to the maximum for each employer, and quarters of social security coverage of the individual). Adjusted gross income from the 100-percent IRS file is then matched to a file of SIPP data. The SIPP cases are then reweighted by controlling to the 1984 IRS controls; that is, a factor f_j, which is the ratio of IRS control in $cell_j$ to the SIPP estimate of persons matched to IRS data with 1984 IRS income in $cell_j$, is applied to persons who fall in $cell_j$ based on the IRS data. Estimates and variances of selected SIPP characteristics will be obtained using the newly created weights and with the weights which do not use this procedure. Fay and Huggins (1988) will provide some indications of the usefulness of this method at the 1988 meetings of the American Statistical Association.

5.5 Sampling for Special Subpopulations

Subgroups of the population are often cited as being more affected by governmental policy than others--the population of persons in poverty, the aged, the Blacks, Hispanics, and participants of Federal income security programs. Early design goals of the ISDP emphasized a concern for improving the reliability of subpopulation estimates. This was exhibited in the emphasis placed in the ISDP on sampling from administrative program lists. Thus, samples were oftentimes drawn from lists of current participants of Federal-or state-administered programs (Kasprzyk, 1983; Bowie and Kasprzyk, 1987).

A Census Bureau Working Group analyzed subsampling (screening) proposals for oversampling special populations. The issue studied concerned the reliability of estimates when different subsampling schemes are introduced. Subsampling characteristics based on income and demographic variables were identified and estimates of reliability for different subsampling rates and characteristics were calculated (U.S. Bureau of the Census, 1985).

This group concluded that subsampling proposals, for a general-purpose income survey like the SIPP, provided only modest gains in precision for low income items and did not outweigh the disadvantages, which included an increase in the complexity of the operation, the loss of a self-weighting design, and large decreases in precision for the middle income items. Because of renewed interest in improving the reliability of the estimates of the "poor" and "near-poor" subpopulations, it is likely that this issue will be reexamined in 1988 and 1989.

6. Response Error

Response error is one aspect of a more general problem, nonsampling error, discussed by Kalton, McMillen, and Kasprzyk (1986) in the context of the SIPP. Response error occurs when incorrect data are recorded on the questionnaire. This can occur for a variety of reasons, such as a faulty questionnaire, memory errors, inappropriate respondents, etc. In this section we briefly describe a response error issue with the SIPP gross flow data and some recent considerations in developing an understanding of the SIPP response error structure.

6.1 SIPP Gross Flow Data

Analysis of program data on a month-to-month basis in ISDP revealed a tendency for reported program turnover to occur between waves of interviewing more often than within the wave (Moore and Kasprzyk, 1984). Analysis using the SIPP data (Burkhead and Coder, 1985) covering month-to-month changes in receipt of income sources for a 12-month period focussed on changes occurring between the last month of one reference period and the first months of the succeeding reference period. The results using SIPP and ISDP data are similar, where an uneven pattern of change is observed, and this pattern is clearly associated with the interviewing scheme. Gross changes are significantly higher between the last month of one reference period and the first month of the next.

Tables 6 and 7 illustrate the magnitude of the problem in the 1984 SIPP Panel. These tables clearly show that respondents report transitions within an interview period differently than between two interview periods. This phenomenon has also been observed in the reporting of the amounts of income received (U.S. Bureau of the Census, 1986c).

Table 6: MONTHLY TRANSITIONS FOR PERSONS: SOCIAL SECURITY AND FOOD STAMPS

(Persons interviewed first 12 months in sample)

		Mean Number of Monthly Transitions within Interview 1	4th month to 5th month	Mean Number of Monthly Transitions within Interview 2	8th month to 9th month	Mean Number of Monthly Transitions within Interview 3
SOCIAL SECURITY	Receiving both months	6,484	6,473	6,651	6,650	6,781
	Not receiving to receiving	56	157	44	134	31
	Receiving to not receiving	19	105	13	73	17
FOOD STAMP	Receiving both months	1,339	1,224	1,365	1,243	1,344
	Not receiving to receiving	72	162	55	140	40
	Receiving to not receiving	53	207	40	190	38

From Burkhead and Coder (1985)

Table 7: MONTHLY TRANSITIONS FOR PERSONS: EARNINGS AND UNEMPLOYMENT COMPENSATION

(Persons interviewed first 12 months in sample)

		Mean Number of Monthly Transitions within Interview 1	4th month to 5th month	Mean Number of Monthly Transitions within Interview 2	8th month to 9th month	Mean Number of Monthly Transitions within Interview 3
EARNINGS	Receiving both months	18,959	18,455	19,158	18,536	19,360
	Not receiving to receiving	676	1118	482	1136	534
	Receiving to not receiving	557	1296	451	1129	453
UNEMPLOYMENT COMPENSATION	Receiving both months	435	313	462	294	343
	Not receiving to receiving	129	234	141	165	90
	Receiving to not receiving	155	241	128	294	105

Adapted from Burkhead and Coder (1985)

The reporting behavior described above is not unique to the SIPP. Hill (1987b) used monthly data from the 1984 and 1985 waves of the Panel Study of Income Dynamics (PSID) to investigate the extent and determinants of excessive change between waves relative to measured change within waves of the PSID. He found that in spite of different question sequences and recall periods, between-wave transitions dominate the within-wave transitions in the PSID just as they do in the SIPP. The main causes for the problem are not known, but questionnaire wording/design, respondent recall error, and the interaction between these two factors seem likely.

Weidman (1986) did an empirical analysis to look for obvious relationships between respondent characteristics and changes in receipt status of a number of income types. He did not detect any sizable relationships between gross change distributions, self/proxy status and nine demographic variables (age, race, sex, education, marital status, household size, tenure, relationship to reference person, and size of metropolitan area) for consecutive months, but did note that more transitions occur when some of the data are imputed. The absence of any notable relationships indicates a need for exploring other ways to understand this problem.

Interest in gross flow estimates remains high. Hubble and Judkins (1986) developed a model to estimate biases in gross flow estimates resulting from response errors, the parameters of which are estimated using SIPP response error rates and the ratios of within-wave and between-wave gross flow estimates. Several strong assumptions, as well as a reinterview program which produces accurate reinterview data on gross flows within the period, are necessary. Weidman (1987) presents linear models that try to represent the relationships between observed and actual transitions. The models are admittedly oversimplified using only survey reported data, but, nevertheless, illustrate the need to obtain more information about the SIPP error structure in reporting receipt of benefits from government transfer programs.

6.2 Recent Considerations in Understanding the SIPP Error Structure

The SIPP program realizes the need to improve understanding of misreporting and misclassification, particularly as they relate to the measurement of flows in income on a month-to-month basis. At a minimum some effort to improve the questionnaire to reduce the problem described in the previous section is necessary. Two types of studies are currently in process: 1) SIPP record check study; and 2) a series of turnover studies. They differ in that the former allows a microlevel comparison of the data, while the latter provides macrolevel comparisons.

SIPP Record Check Study

One way to address the SIPP error structure in reporting receipt of program benefits and amounts is to develop validation studies of items common to both survey records and administrative records. The SIPP program has initiated such a study to investigate response quality issues.

The goal is the improved understanding of the quality of the SIPP data and, ultimately, the development of quantitative estimates of response and non-response errors in order to adjust the survey data or modify survey procedures to obtain better quality data. The research questions addressed in this study

include: 1) the quality of the respondent reports of receipt of program benefits for a variety of state and Federally administered transfer programs; 2) the quality of benefit dollar amount reporting for these programs; 3) demographic correlates of report quality; 4) extent of misclassification errors; 5) the effects of self-proxy respondent status on report quality; and 6) between-wave recipiency turnover effects. Four state-administered programs and six Federally administered programs are included in the study. Moore and Marquis (1987) provide very preliminary results, suggesting that reporting problems are different for the Aid to Families with Dependent Children (AFDC) and the food stamp programs--the former having a net underreporting and a time-placement problem for reporting a transition in program status, while the latter has only a time-placement problem. Moore and Marquis (1988) will provide further results at the 1988 meeting of the American Statistical Association.

Macrolevel Turnover Studies

Singh, Weidman, and Shapiro (1988) summarize research on the measurement of transitions in the SIPP. They describe several studies which compare aggregate statistics from administrative data with transition rates as measured in the SIPP. With regard to the food stamp program, they showed that SIPP transition rates are very similar to those observed from an administrative record sample. A comparison of SIPP to AFDC administrative data showed that average start-up and average exit rates were lower in the SIPP though not statistically significant. Using aggregate data from the Social Security Administration, Singh, et al.(1988) report that SIPP measured significantly higher start-up and exit rates for the Supplemental Security Income Program. These studies are limited, but are, nevertheless, useful in trying to roughly assess the magnitude of bias in estimates of transitions.

Other research to reduce the microlevel gross flow reporting problem is planned:

- providing more information on this problem with interviewer training materials;
- placing significantly more emphasis on data quality during interviewer training;
- analyzing the effectiveness of the changes instituted to reduce the problem in the 1988 Panel questionnaires;
- developing a calendar as a data collection aid to assist the respondent in recalling the timing of certain transitions;
- analyzing existing data to study interviewer effects on the proportion of between-wave transitions;
- analyzing self-proxy response status and their effect on the measurement of transition;
- conducting exploratory research in a cognitive laboratory setting to generate hypotheses/models for improving the measurements of transitions;
- planning the development of an alternative data collection method, such as a time-line calendar.

This work will occur during 1988 and 1989.

7. Conclusion

The SIPP program has been ambitious from its inception and expectations for the program have been high. Morton Hunt in his Profiles of Social Research referred to the SIPP as "the most exciting thing going on in Social Science in the 1980's." Enthusiasm and high expectations have continued to characterize the program. This paper has described the program and tried to show the nature of the research undertaken under its aegis. As such, it has described the principal research issues of the program from its earliest days to the present time. Recently more effort has been expended on the "evaluation" and analysis of the data collected in the SIPP. Vaughan (1988), Coder (1988), Farley and Neidert (1988), King, Petroni, and Singh (1987), and Singh, Weidman, and Shapiro (1988) describe how the data compare to other established data sets. The latter two research papers also describe the sources of nonsampling error and the magnitude of sampling error in the SIPP. Numerous presentations of analytic results from the SIPP data have been made at meetings of demographers, sociologists, economists, and statisticians. Approximately 60 analytic research papers will be presented in 1988.

The current activity suggests an acceptance of the new data set; however, further methodological research is still in order. Time-in-sample bias, a source of nonsampling error in all panel surveys, has not been investigated in the context of the SIPP. The combining of data from two SIPP panels, an integral part of the survey design, has not yet been undertaken at this time. Indeed, estimates from two or more panels must be rigorously compared. Because of the cumulative nonresponse rates, additional research to improve nonresponse compensation procedures is desirable. Finally, the availability of longitudinal microdata files allows the possibility of analyzing the data longitudinally. Analyses conducted will lead to questions and investigations into edits, imputations, and estimation procedures used on the longitudinal products.

As in all large-scale, continuing survey efforts, research is needed to improve understanding of the effects of survey methods on the data collected. A relatively new survey, like the SIPP, which is complex in its implementation requires a commitment to understanding the measurement process. The wide range of topics discussed above--questionnaire design, data collection, longitudinal concepts and estimation, and response error--illustrate where the interest and emphasis was placed during the development program and the first few years of the SIPP program. One hopes that the program's level of commitment to research methods can be maintained in the future.

Acknowledgments

The author gratefully acknowledges the secretarial and clerical help of Ms. Hazel Beaton and Miss Debra Ganni in the preparation of this manuscript.

REFERENCES

BOWIE, C. and D. KASPRZYK (1987).
"A Review of the Use of Administrative Records in the Survey of Income and Program Participation," SIPP Working Paper Series, No. 8721. Washington, D.C.: U.S. Bureau of the Census.

BURKHEAD, D. and J. F. CODER (1985).
"Gross Changes in Income Recipiency From the Survey of Income and Program Participation," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 351-356.

CARMODY, T. J., D. P. FISCHER, and A. J. MEIER (1988).
"Survey of Income and Program Participation Experiments and Their Results." Paper presented to the Section on Survey Research Methods, American Statistical Association, August 1988.

CHAKRABARTY, R. P. (1986).
"Composite Estimation for SIPP: A Preliminary Report," SIPP Working Paper Series, No. 8610. Washington, D.C.: U.S. Bureau of the Census.

CHAPMAN, D. W., L. BAILEY, and D. KASPRZYK (1986).
"Nonresponse Adjustment Procedures at the U.S. Bureau of the Census," Survey Methodology, 12(2), pp. 161-180.

CITRO, C. F. (1985).
"Alternative Definitions of Longitudinal Households in the Income Survey Development Program: Implications for Annual Statistics," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 381-386.

CITRO, C.F., D. J. HERNANDEZ, and R. A. HERRIOT (1986).
"Longitudinal Household Concepts in SIPP: Preliminary Results," Proceedings of the U.S. Bureau of the Census Second Annual Research Conference, pp. 598-619.

CITRO, C.F., D. J. HERNANDEZ, and J. E. MOORMAN (1986).
"Longitudinal Household Concepts in SIPP," Proceedings of the Social Statistics Section, American Statistical Association, pp. 361-366.

CODER, J. F. (1980).
"Some Results From the 1979 Income Survey Development Program Research Panel," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 540-545.

CODER, J. F. (1988).
"Exploring the Quality of Data From the 1984 SIPP Panel File." Paper presented at the SIPP/SSRC Conference on "Individuals and Families in Transition: Understanding Change Through Longitudinal Data," in Annapolis, Maryland.

CODER, J. F., and A. M. FELDMAN (1984).
"Early Indications of Item Nonresponse in SIPP," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 693-697.

ERNST, L. R., (1988).
"Weighting Issues for Longitudinal Household and Family Estimates," Panel Surveys (Eds. D. Kasprzyk, G. J. Duncan, G. Kalton, M. P. Singh). New York: John Wiley and Sons, Inc., forthcoming.

ERNST, L. R., D. L. HUBBLE, and D. R. JUDKINS (1984).
"Longitudinal Family and Household Estimation in SIPP," Proceedings of the Section on Survey Research Methods, American Statistical Association pp. 670-676.

FARLEY, R. and L. J. NEIDERT (1988).
"Analyzing the Characteristics of Blacks: A Comparison of Data From the SIPP and the CPS." Paper presented to the Social Statistics Section, American Statistical Association, August 1988.

FAY, R. E. and V. J. HUGGINS (1988).
"Use of Administrative Data in SIPP Longitudinal Estimation." Paper presented to the Section on Survey Research Methods, American Statistical Association, August 1988.

GBUR, P. and S. DURANT (1987).
"Testing Telephone Interviewing in the Survey of Income and Program Participation and Some Early Results." Paper presented at the International Conference on Telephone Survey Methodology, in Charlotte, North Carolina.

GRIFFITH, J. E. (1978).
"Impact of a Longitudinal Survey on the Data Base," ISDP Working Paper No. 1. Washington, D.C.: U.S. Department of Health and Human Services. To obtain contact SIPP ACCESS, Institute for Research on Poverty, University of Wisconsin, 1180 Observatory Drive, Madison, WI 53706 (608-262-6358).

HEERINGA, S. G., and J. M. LEPKOWSKI (1986).
"Longitudinal Imputation for the SIPP," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 206-210.

HERRIOT, R. A. (1983).
"The Use of Administrative Records in Social and Demographic Statistics." Paper presented at the Meeting of the International Statistics Institute in Madrid, Spain.

HILL, D. H. (1987a).
"Response Errors in Labor Surveys: Comparisons of Self and Proxy," Proceedings of the Bureau of the Census Third Annual Research Conference, pp. 299-319.

HILL, D. H. (1987b).
"Response Errors Around the Seam: Analysis of Change in a P. l with Overlapping Reference Records," Proceedings of the Section on Su y Research Methods, American Statistical Association, pp. 210-215.

HUANG, H. (1984).
"Obtaining a Cross-Sectional Estimate From a Longitudinal Sur : Experience of the ISDP," Proceedings of the Section on Survey Research Me nods, American Statistical Association, pp. 670-675.

HUBBLE, D. L. and D. R. JUDKINS (1986).
"Measuring the Bias in Gross Flow Estimates in the Presence of Auto-Correlated Response Errors," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 237-242.

HUGGINS, V. J. (1987a).
"Evaluation of Missing Wave Data From the Survey of Income and Program Participation," Proceedings of the Section on Survey Research Methods, American Statistical Association, Washington, D.C., pp. 205-209.

HUGGINS, V. J., M. E. SAMUHEL, and L. WEIDMAN (1985).
"Methodologies for Imputing Longitudinal Survey Items," SIPP Working Paper Series No. 8506. Washington, D.C.: U.S. Bureau of the Census.

HUGGINS, V. J. and L. WEIDMAN (1986a).
"An Investigation of Model-Based Imputation Procedures Using Data From the Income Survey Development Program," SIPP Working Paper Series No. 8603. Washington, D.C.: U.S. Bureau of the Census.

HUGGINS, V. J. and L. WEIDMAN (1986b).
"An Investigation of the Imputation of Monthly Earnings for the Survey of Income and Program Participation Using Regression Models," SIPP Working Paper Series No. 8607. Washington, D.C.: U.S. Bureau of the Census.

HUNT, M. (1986).
"Chapter 3: Sampling Social Reality: A Complex New Survey Measures the Impact of the Government's Social Programs on the American Family," Profiles of Social Research: The Scientific Study of Human Interactions. New York: Russell Sage Foundation, pp. 98-154.

JEAN, A. C. and E. K. McARTHUR (1984).
"Some Data Collection Issues for Panel Surveys With Application to the Survey of Income and Program Participation," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 745-750.

JEAN, A. C. and E. K. MCARTHUR (1987).
"Tracking Persons Over Time," SIPP Working Paper Series No. 8701. Washington, D.C.: U.S. Bureau of the Census.

JUDKINS, D. R., D. L. HUBBLE, R. P. SINGH, D. B. McMILLEN, and L. R. ERNST (1984).
"Weighting of Persons for SIPP Longitudinal Tabulations," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 676-681.

KALTON, G. (1986).
"Handling Wave Nonresponse in Panel Surveys," Journal of Official Statistics, 2(3), pp. 303-314.

KALTON, G., D. KASPRZYK, and D. B. MCMILLEN (1988).
"Nonsampling Errors in Panel Surveys," Panel Surveys (Eds. D. Kasprzyk, G. J. Duncan, G. Kalton, and M. P. Singh). New York: John Wiley & Sons, Inc., forthcoming.

KALTON, G. and J. M. LEPKOWSKI (1983).
"Cross-wave Imputation," Technical, Conceptual, and Administrative Lessons of the ISDP (Ed. M. H. David). New York: Social Science Research Council, pp. 171-198.

KALTON, G. and J. M. LEPKOWSKI (1985).
"Following Rules in SIPP," Journal of Economic and Social Measurement, 13 (3/4), pp. 319-329.

KALTON, G., J. M. LEPKOWSKI, and T. K. LIN (1985).
"Compensating for Wave Nonresponse in the 1979 ISDP Research Panel," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 372-377.

KALTON, G., D. B. McMILLEN, and D. KASPRZYK (1986).
"Nonsampling Error Issues in the Survey of Income and Program Participation," Proceedings of the U.S. Bureau of the Census Second Annual Research Conference, pp. 147-165.

KALTON, G. and M. E. MILLER (1986).
"Effects of Adjustments for Wave Nonresponse On Panel Survey Estimates," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 194-199.

KASPRZYK, D. (1983).
"Social Security Number Reporting, the Use of Administrative Records, and the Multiple Frame Design in the Income Survey Development Program," Technical, Conceptual, and Administrative Lessions of the Income Survey Development Program (ISDP), (Ed. M. H. David). New York: Social Science Research Council, pp. 171-198.

KASPRZYK, D. and G. KALTON (1983).
"Longitudinal Weighting in the ISDP," Technical, Conceptual, and Administrative Lessons of the ISDP (ed. M. H. David). New York: Social Science Research Council, pp. 155-170.

KASPRZYK, D. and D. B. McMILLEN (1987).
"SIPP: Characteristics of the 1984 Panel," Proceedings of the Social Statistics Section, American Statistical Association, pp. 181-186.

KING, K. E., R. J. PETRONI, and R. P. SINGH (1987).
"Quality Profile for the SIPP," SIPP Working Paper Series No. 8708. Washington, D.C.: U.S. Department of Commerce, Bureau of the Census.

KOBILARCIK, E. and R. P. SINGH (1986).
"SIPP Longitudinal Estimation for Persons' Characteristics," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 214-219.

LAMAS, E. J. and J. M. McNeil (1984).
"The Measurement of Household Wealth in the SIPP," Proceedings of the Social Statistics Section, American Statistical Association, pp. 484-489.

LAMAS, E. J. and J. M. McNEIL (1987).
"An Analysis of the SIPP Asset and Liability Feedback Experiment," Proceedings of the Social Statistics Section, American Statistical Association, pp. 195-199.

LANE, J. P. (1978).
"Four Alternative Ways of Defining "Annual" Analysis Units," ISDP Working Paper No. 6. Washington, D.C.: U.S. Department of Health and Human Services. To obtain contact SIPP ACCESS, Institute for Research on Poverty, University of Wisconsin, 1180 Observatory Drive, Madison, WI 53706 (608-262-6358).

LEPKOWSKI, J. M. (1988).
"The Treatment of Wave Nonresponse in Panel Surveys," Panel Surveys (Eds. D. Kasprzyk, G. J. Duncan, G. Kalton, and M. P. Singh). New York: John Wiley and Sons, Inc., forthcoming.

McMILLEN, D. B. and R. A. HERRIOT (1985).
"Toward a Longitudinal Definition of Households," Journal of Economic and Social Measurement, 13 (3/4), 349-360.

McMILLEN, D. B. and D. KASPRZYK (1985).
"Item Nonresponse in the SIPP," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 360-365.

MOORE, J. C. and D. KASPRZYK (1984).
"Month-to-Month Recipiency Turnover in the ISDP," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 726-731.

MOORE, J. C. and K. H. MARQUIS (1987).
"Using Administrative Data to Evaluate the Quality of Survey Estimates," SIPP Working Paper Series No. 8806, Washington, D.C.: U.S. Bureau of the Census.

MOORE, J. C. and K. H. MARQUIS (1988).
"Using Administrative Record Data to Describe SIPP Response Errors." Paper presented to the Section on Survey Research Methods, American Statistical Association, August 1988.

NELSON, D., C. BOWIE, and A. WALKER (1987).
"SIPP Sample Loss and the Efforts to Reduce It," Proceedings of the Bureau of the Census Third Annual Research Conference, pp. 629-643.

OLSON, J. A., (Ed.) (1980)
ISDP Reports from the Site Research Test. Office of the Assistant Secretary for Planning and Evaluation. Washington, D.C.: U.S. Department of Health and Human Services. To obtain contact SIPP ACCESS, Institute for Research on Poverty, University of Wisconsin, 1180 Observatory Drive, Madison, WI 53706 (608-262-6358).

PETRONI, R. J., and K. E. KING (1988).
"Evaluation of the Survey of Income and Program Participation's Cross-Sectional Nonintervew Adjustment Method." Paper presented to the Section on Survey Research Methods, American Statistical Association, August 1988.

ROMAN, A. M. and D. V. O'BRIEN (1984).
"The Student Follow-up Investigation of the 1979 ISDP," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 732-737.

SAMUHEL, M. E. and V. J. HUGGINS (1984).
"Longitudinal Item Imputation in a Complex Survey," Proceedings of the Section on Survey Research Methods, American Statistical Association. pp. 688-692.

SCHEUREN, F. (1983).
"Design and Estimation for Large Federal Surveys Using Administrative Records," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 288-293.

SINGH, R. P., L. WEIDMAN, and G. M. SHAPIRO (1988).
"Quality of the SIPP Estimates." Paper presented at the SIPP/SSRC Conference on "Individuals and Families in Transition: Understanding Change Through Longitudinal Data" Annapolis, Maryland.

U.S. BUREAU OF THE CENSUS (1985).
"SIPP Research on SIPP Oversampling/Subsampling." Memorandum from R. P. Singh to G. M. Shapiro and D. Kasprzyk, Population Division (August 12, 1985).

U.S. BUREAU OF THE CENSUS (1986a).
"Additional Results from the SIPP Telephone Test." Memorandum from J. F. Coder to A. Norton, Population Division (April 9, 1986).

U.S. BUREAU OF THE CENSUS (1986b).
"SIPP: Recall Effect Study." Memorandum from R. P. Singh to G. M. Shapiro, Statistical Methods Division (October 2, 1986).

U.S. BUREAU OF THE CENSUS (1986c).
"Monthly Transitions From the SIPP Longitudinal Research File." U.S. Bureau of the Census Memorandum from J. F. Coder to P. Schneider, Population Division (May 20, 1986).

U.S. Bureau of the Census (1987).
"Research Plans." Memorandum for the Record from V. J. Huggins, Statistical Methods Division (April 13, 1987).

U.S. BUREAU OF THE CENSUS (1988a).
"Comparisons of Alternative Annual Estimates of Wage and Salary Income from SIPP." Memorandum from J. F. Coder to G. W. Green, Population Division (March 29, 1988).

U.S. BUREAU OF THE CENSUS (1988b).
"SIPP 84: Specification for Panel File Longitudinal Weighting of Persons." Memorandum from P.J. [illegible]e to T. Walsh, Demographic Surveys Division (November 1988).

U.S. BUREAU OF THE CENSUS (1988c).
"Measurement of the Attrition through the Completed SIPP 1984 Panel: Preliminary Results." Memorandum from E. K. McArthur to D. Kasprzyk, Office of the Director (March 4, 1988).

VAUGHAN, D. R. (1988).
"Reflections on Income Estimates From the Initial Panel of the SIPP." Paper presented at the SIPP/SSRC Conference on "Individuals and Families in Transition: Understanding Change Through Longitudinal Data" in Annapolis, Maryland.

WEIDMAN, L. (1986).
"Investigation of Gross Changes in Income Recipiency From the Survey of Income and Program Participation," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 231-236.

WEIDMAN, L. (1987).
"Examination of Relationships Between Actual and Reported Changes in the SIPP," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 216-220.

WEIDMAN, L., K. E. KING and T. WILLIAMS (1988).
"Evaluation of the SIPP Asset Feedback Experiment." Paper presented to the Section on Survey Research Methods, American Statistical Association, August 1988.

WHITE, G. D. and H. HUANG (1982).
"Mover Follow-up Costs for the Income Survey Development Program," Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 376-381.

YCAS, M. A. and C. A. LININGER (1981).
"The Income Survey Development Program: Design Features and Initial Findings," Social Security Bulletin, 44 (11), Baltimore, Md.: Social Security Administration, pp. 13-19.

Survey of Income and Program Participation

Census Bureau Microdata: Providing
Useful Research Data While Protecting
The Anonymity of Respondents

8829 76

Gerald W. Gates
Bureau of the Census

November 1988

This paper was prepared for presentation at the annual meeting of the American Statistical Association in New Orleans, August 22-25, 1988. The views expressed are the author(s) and do not necessarily reflect those of the Census Bureau.

TABLES OF CONTENTS

INTRODUCTION..... 1

LEGAL CONSIDERATIONS..... 4

PUBLIC USE SOLUTIONS..... 4

The Microdata Review Process..... 4

Research on Microdata Disclosure Risk and Reduction..... 6

Public Use Alternatives to Microdata..... 8

ADMINISTRATIVE SOLUTIONS..... 10

Non-Title 13 Surveys..... 11

Use of Special Sworn Employees..... 12

LEGAL OPTIONS..... 15

CONCLUSIONS..... 18

REFERENCES

Table 1. Objectives Of The ASA/NSF/Census Research Program.....22-23

Table 2. Census Bureau Regional Offices..... 24

INTRODUCTION

The U. S. Census Bureau has provided public use microdata as a component of its decennial census data products since 1963 when we released a one-in-one-thousand sample file for the 1960 Decennial Census. Since then, microdata files have become an integral part of our decennial census and demographic surveys programs. As a result, researchers in other Government agencies and research institutes have been able to conduct important policy and planning studies that could not be answered through the use of published tabulations. Were it not for public use microdata files, the only way these studies could be done, if at all, would be by contracting with the Census Bureau for special tabulations. This is not the preferred solution for several reasons. First, these special requests are totally dependent on programming and computer support that is committed to routine Census work. Therefore, the time required to complete the work does not always satisfy user needs. Second, statistical analyses do not always turn out the way researchers intended. They may want to change the variables or the analytical methods after they see the initial results. Finally, in contrast to the costs of using available staff and micro-computers, the costs of using Census Bureau main-frame computers and programmers may exceed the available resources for the project.

The advent of public use files has eliminated many of these problems but has introduced some new ones both for the researcher and for the Census Bureau. Because of the flexibility available when using microdata files, the broad access to high speed computers, and the increased sophistication of data users, there has been an increased use of this medium in the 1970's and, particularly in the 1980's. With this increased use has come increased demand for detailed information that was excluded from or curtailed on public use files to protect the identity of survey and census respondents. The statute (Title 13) under which the Census Bureau operates requires that when we collect and publish data under this authority we not publish results that can be used to identify a particular respondent. Realizing that it may not be possible to release data from which it is absolutely impossible to identify an individual, we strive to ensure that the risk that the data will be used to identify someone on the file is extremely small. For example, current microdata disclosure protection criteria prevent the release of geographic identifiers for areas with small populations, extreme values for continuous variables, and information that is obtained from

---

This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributable to the author and do not necessarily reflect those of the Census Bureau.

or matchable to administrative records systems. These restrictions prevent survey sponsors from conducting some analyses, such as certain microsimulations; reprocessing the individual responses; or having their own administrative data appended to the survey results. They also inhibit the potential of surveys we sponsor (for example, the Survey of Income and Program Participation (SIPP)) for program policy research by other Government agencies.

Here are some recent examples of requests for demographic microdata that could not be satisfied because of confidentiality concerns:

- The General Accounting Office requests a file linking SIPP data to Social Security beneficiary records. This file is needed for a study related to a disparity in Social Security benefits between adjacent cohorts of retirees. The information from the Social Security records are match keys that could be used by the SSA to identify SIPP respondents.

- The Economic Research Service of the Department of Agriculture wants a file showing non-metro status of SIPP respondents in order to assess the economic well-being of non-metro residents in terms of their wealth, asset holdings, and participation in Government programs. These non-metro designations, in combination with the geography on the released public use files, reveal areas of fewer than 100,000 persons.

- The National Opinion Research Center (NORC) requests a special 1980 census public use file with records linked to tract and SMSA data. This study, linking people to their immediate neighborhoods (tracts) and the larger area in which they live (SMSA), is part of a three year study of racial segregation in the U.S.. Tracts and some SMSAs contain populations of fewer than 100,000 persons.

- Princeton University requests exact date of birth on a SIPP microdata tape in order to research the Selective Service draft lotteries held in the U.S. in the 1970s. (Lottery numbers were assigned to young men based on birth date.) Since date of birth is available on many administrative records files, it is an excellent match key and an additional risk to identifying SIPP respondents.

- The National Institute on Aging (NIA) wants to conduct a followup interview with respondents to the Longitudinal Retirement History Survey conducted in the 1970s. One condition for funding a followup survey is that a microdata file be made available for research studies supported by the NIA. Such a file would be potentially matchable to administrative records information maintained by the Social Security Administration.

- The NORC would like SSA earnings history data added to a SIPP microdata file to be used as a control group in an evaluation of the Job Training Partnership Act (JPTA) manpower training system. Data for the control group would be used to measure the impact on outcomes such as earnings, labor force participation patterns, and welfare recipiency of the JTPA program relative to a population of non-participants.

- The Bureau of Labor Statistics would like access to finer geography and certain longitudinal matching variables on a Current Population Survey (CPS) public use file. This survey is sponsored jointly by the BLS and the Census Bureau. BLS wants this additional detail in order to conduct statistical research, facilitate longitudinal analysis of the data, and develop small area estimates.

Users of data from the Census Bureau economic surveys and censuses have a more basic problem when it comes to microdata. Namely, the Census Bureau has not released microdata on businesses because of the unique visibility of establishments, the availability of private sector data bases, and the effects such files would have on our ability to produce subsequent special purpose tabulations. Nevertheless, demand continues to grow for public use files on businesses; particularly those relating to the manufacturing sector. For example, the Census Bureau has developed a longitudinal file of manufacturers called the Longitudinal Establishment Data (LED) file. In a conference sponsored by the Census Bureau in 1984, more than 100 economists interested in the LED expressed their desire for a public use LED file. The only alternative they saw--submitting special requests for analyses to the Census Bureau--was totally unacceptable because of the limited utility of releasable products and the timing and cost factors, (Govoni-Waite, 1985).

Aside from the interests of our users, the Bureau of the Census must also be concerned about whether the protections afforded these public use files are sufficient. While high speed computers have made public use files more attractive, they have also increased public concern about potential abuses to individual privacy resulting from the creation of large integrated databases. In recent years, events in West Germany, Sweden and other European countries regarding government databases have highlighted this concern, (Butz, 1985; Dalenius, 1988). Moreover, widely publicized exploits of computer hackers have raised fears that, given enough patience, someone could defeat any scheme designed to protect confidentiality. On top of this we know very little about the true risk of someone identifying a respondent on a public use microdata file. Statisticians are just now beginning to quantify the disclosure risks associated with microdata, (Duncan-Lambert, 1987; Paass, 1985; Spruill, 1983). Perhaps instead of seeking ways to provide more detailed public use microdata we should be looking for alternatives that contain fewer unknowns.

With the growing demand for microdata products that cannot be made public under current guidelines and the lack of an acceptable quantitative measure of disclosure risk, the Bureau has undertaken to find solutions that provide our users with the data they want and our respondents with the data protection assurances they are entitled to. This paper describes our current plans in terms of public use microdata, publicly releasable alternatives to microdata, and administrative arrangements. I describe some applications of these solutions to recent requests. Finally, I discuss legal arrangements that have been recommended as ways of extending the obligation for protecting confidentiality to the users of microdata.

LEGAL CONSIDERATIONS

Release of individual data by the Census Bureau is restricted by Title 13, United States Code. Only sworn officers and employees of the Census Bureau are allowed to examine individual reports furnished under the provisions of this title. As needed, we have the authority by Section 23 to "utilize temporary staff, including employees of Federal, State, or local agencies or instrumentalities, and employees of private organizations to assist (us) in performing the work authorized by this title, but only if such temporary staff is sworn to observe the limitations imposed by Section 9 of this title." Section 9 (a) states that the Census Bureau may not "use the information furnished under the provisions of this title for any purpose other than the statistical purposes for which it is supplied" and may not "make any publication whereby the data furnished by any particular establishment or individual under this title can be identified."

PUBLIC USE SOLUTIONS

Public use microdata are data products the Census Bureau releases for general, unrestricted statistical and nonstatistical use. As a result of our legal requirements, we must ensure that any microdata product we release to the public is anonymous (no individual identifiers) and that it will remain so. Consequently, individual characteristics on the file must be evaluated to determine if they can be employed to uniquely describe an individual in the population from which the sample was selected. This evaluation procedure involves making some assumptions about what external information is available, whether it is accessible, and the amount of effort required to retrieve it. Where records are not available, we consider the visibility of persons (that is, things about them that are public knowledge and would be revealed in the file).

The Microdata Review Process

Prior to 1981, the Census Bureau's microdata disclosure reduction criterion consisted of a 250,000 minimum requirement for the population residing in sample areas that represent the finest geographic area to be shown on the file. Additional disclosure

reduction measures were established on a case-by-case basis by the Census Bureau staff responsible for releasing the file. In 1981, other criteria were established, including a new population minimum of 100,000 within sample areas; although a higher minimum could be set if the nature of the file warranted greater restrictions. At this time the Census Bureau also created a Microdata Review Panel (MRP) to review and approve all microdata files prior to release.

The Panel's membership included staff representing the Directorates for Statistical Standards and Methodology, Economic Programs, and Demographic Programs; and the Data Users Services Division and the Program and Policy Development Office. The MRP was given broad authority to require additional masking techniques to reduce disclosure risk. These include data grouping or aggregation, addition of random noise, rounding responses, and in some cases, suppression. In order to allow for a smooth transition and minimize the disruption to current microdata users, files that were released prior to 1981 were not recalled and surveys that were currently in the field were not subject to MRP review. Continuing surveys come under MRP review only after the sample is redesigned, the content of the questionnaire is materially changed or the content of the file is expanded.

A typical microdata review consists of the following steps:

1. The sponsoring Census Bureau division submits a formal request to release a file. This request includes:

   - tables showing population counts in identifiable geographic areas;
   - a description of the survey design, sampling procedures, and weighting scheme;
   - a checklist identifying potential disclosure problems with the file, including the existence of external files (e.g. administrative records) which contain data items similar to the proposed release;
   - proposed solutions to these disclosure problems including topcodes, recodes, and deletions; and
   - a data dictionary or annotated questionnaire for the proposed file indicating which items are to be recoded, topcoded, grouped or suppressed.

2. The MRP meets to review the request taking into consideration:

   - Disclosure reduction requirements imposed on previous releases (if any) from the subject survey.

   - If the survey is longitudinal, whether the proposed geography has been changed from the previous release? If it has, could the current and previous releases be matched on characteristics to reveal areas of fewer than 100,000 persons?
   - What information from the proposed file is available from

external files; including those available to the survey sponsor?
- If the survey sample was drawn from other Census Bureau surveys or censuses, were microdata files released from those programs and what information did they contain?
- The uniqueness and degree of visibility of characteristics on the file in conjunction with the proposed geography (for example, residence in a particular institution).

3. The MRP approves the file for release as proposed; requires specific modifications; suggests possible solutions that the division/sponsor may accept or propose an alternative; or rules that a microdata product is not possible given the requirements of the sponsor.

The decisions of the MRP are partly subjective in that no quantitative measures of disclosure risk are available for each file. The panel members varied backgrounds within the Census Bureau tend to promote a balance in the review process which recognizes the needs of our users while emphasizing our obligations to respondents. In recent years, with increased demands for more detailed geography and administrative data appended to surveys, the Panel's seemingly conservative stance has come under criticism by users.

Research on Microdata Disclosure Risk and Reduction

In order to provide a more scientific approach to evaluating microdata disclosure risk, the Census Bureau has established a permanent staff to conduct research on disclosure risk measurement and reduction, (Greenberg, 1988). This Census Bureau Confidentiality Staff is currently undertaking "reidentification studies" for the Survey of Income and Program Participation and the 1990 Decennial Census sample files. These studies involve measuring (or quantifying) the risk of disclosure (identification of a respondent) and designing methods to reduce this risk. Reidentification studies for the proposed decennial census microdata files will be done using the 1980 Decennial Census five-percent public use microdata file and the entire 1980 Census file. The files will be matched using rules that incorporate knowledge of what information is available on external files. The SIPP study involves a similar investigation with a special focus on the effect of geographic detail on levels of disclosure risk.

A logical extension of this research is a methodological evaluation of various masking techniques. In the early planning stages, this work would involve designing methods to evaluate and optimize the effectiveness of various techniques with respect to reducing disclosure risk and maintaining the statistical utility of the data. The schemes we will look at include: 1) recoding responses into intervals; 2) rounding responses; 3) recoding responses into categories; and 4) adding random noise to the responses. We will evaluate the effectiveness of these techni-

ques to reduce disclosure risk and incorporate them, as necessary, depending on the results obtained in the study of disclosure risk, (Greenberg, 1988).

Some Applications to Demographic Microdata

There have been a few instances where we have developed special purpose masking schemes which involve the introduction of random noise. One case involved a microdata file from the Continuous Longitudinal Manpower Survey (CLMS) which we conduct for the Department of Labor to evaluate the effectiveness of the Comprehensive Employment and Training Act (CETA) of 1973. The public use files from this survey contain earnings data matched from SSA administrative records. Since this survey was in effect prior to 1981, the microdata files had not come under MRP review and had not been subject to the systematic analysis of risks involved with files linked to administrative records. Through the addition of random noise and data transformation, we were able to continue to release public use files that adequately protected the confidentiality of respondents, (Kim, 1986). However, we were not able to provide the full range of income data through these techniques.

On occasion, we have developed masking schemes in response to user requests for special purpose data files. An example is the previously mentioned request from the NORC for census tract characteristics on a 1980 census sample file. The Census Bureau Confidentiality Staff has developed a two part approach to this problem. First, they are developing variance-covariance matrices of the data, along with the means, based on the modeling that NORC has planned (see "Public Use Alternatives to Microdata" below). Also, we will prepare a microdata file containing tract characteristics to which noise has been added in order to reduce the risk of tract identification. This approach was developed in consultation with the NORC who determined that the noise would not unduly affect the utility of the data.

Potential Applications to Economic Microdata

The Census Bureau has recently explored the utility of surrogate public use files, involving data transformations, as a means of releasing sensitive economic microdata. To be useful, these transformed files must preserve the correct estimates of the true economic model; allow the analysis of subsets of the data cross-sectionally and longitudinally; and allow expansion of the file to include new economic variables and a link to outside sources, (McGukin-Nguyen 1988). Two types of transformations have been suggested: 1) stochastic transformations which involve adding random noise to the original data while preserving the mean and variance of the variables and the covariance relationships between variables (Kim 1986); and 2) non-stochastic transformations which provide for the release of the data in ratio form, (Monahan, 1986). Each of these methods has merit but each has limitations with respect to the types of economic research for which it will provide a suitable database. McGukin and Nguyen have described the disclosure issues involved in each of these

types of surrogate files and the usefulness of transformation techniques in providing correct estimators for a particular class of single-equation economic models. They conclude that:

> It is extremely important to develop precise criteria for evaluating the disclosure risk. Without such criteria, evaluating a microdata public use file in terms of disclosure is almost impossible. But, we emphasize that disclosure free files are not enough. Such files must be useful and we think the best hope for developing a public use file lies in focusing on surrogate data files which allow researchers to estimate common economic models. Finally, because current economic analysis often uses multi-equation economic models, further research into transformation techniques should take into account these models as well.

Public Use Alternatives to Microdata

There are occasions where traditional masking techniques do not allow for the release of microlevel information needed by policy makers concerned with both economic and social programs. In some cases the sensitivity of the data (for example, information on businesses) or the amount of masking required will prohibit the release of a useful microdata file. That is, the masking necessary to protect the file will destroy important relationships among the variables in the file. To handle these situations, we are experimenting with the release of data tapes containing summary statistics. In addition, we are considering the development of test files as a means of allowing researchers to interact with the internal microdata without having direct access to the files.

Tabulations of Summary Statistics

One category of products would include tabulations of summary statistics, such as microaggregations, whereby individual records are grouped according to specified criterion variables and responses are replaced with averages for the group, (Wolf, 1988). This approach, which is operationally straight-forward, has been suggested as a way to provide access to economic microlevel information, (Govoni-Waite, 1985). It is not a panacea, however, since certain useful properties of the individual data will be lost. One major area for investigation in this approach is to determine rules for grouping establishments. Some users will not be satisfied with the rules that are chosen and this inflexibility is a major limitation to this approach.

Another summary statistic approach we are considering for more general application is the release of variance-covariance or correlation matrices of the data, (McGukin-Nguyen, 1988). Such files allow the outside user to obtain information needed for

producing linear regression estimates based on the underlying microdata and provide excellent confidentiality protection since any given covariance matrix can derive from an infinite number of data sets. As with all summary statistics, the biggest disadvantage with correlation matrices is that they are relatively inflexible for general statistical use. Different users will require different matrices just as the same user may require new columns in his matrix as the analysis proceeds.

Remote Access or Test File Approach

Another public use alternative which we are considering resembles a procedure used by the Luxembourg Income Study (LIS) to provide worldwide access to the LIS database through a telecommunications network, (Rainwater-Smeeding, 1988). In the case of the LIS, certain databases were loaned to the Study by countries with severe privacy and confidentiality restrictions. Since no public use files were permitted, and due to the cost and inconvenience of traveling to Luxembourg to work with the database, an alternative had to be developed.

The solution involved the use of an electronic file transfer network over which users submit program jobs to be run by LIS staff on the database housed in Luxembourg. This process depends entirely upon a user package created by the LIS staff containing: 1) a technical description of the data file; 2) a description of the variables for each country's file including summary statistics; 3) a codebook; 4) recodes for income definitions; 5) a sample data file containing 200 records from each country; and 6) information on available software packages. With this package, the potential user can plan a study, program tabulations, and determine, to some degree, the utility of products created using the "live" data.

Important considerations for the data provider are 1) the degree of confidentiality protection afforded the test file; 2) the physical separation of the users from the live data through the intervention of the LIS staff; and 3) confidentiality measures applied to the tabular output. Important for the user are: 1) familiarity with required software (SPSSX); 2) the degree to which the test file resembles the complete data file and 3) the time required from submission of jobs to the receipt of output. Regarding the test file, LIS provides live records, without personal identifiers, that contain little or no geographic detail but no additional masking. In the absence of public use files containing geographic identifiers, these records should be relatively anonymous. Jobs that are received are held until released by LIS staff. Once submitted, LIS software checks the programs for consistency. Completed jobs are checked by other software for minimum cell size and to ensure that the individual records are not being transmitted. Turn-around time is not instantaneous but, given that nearly everything is automated, it can generally be measured in hours rather than days.

The application of this approach at the Census Bureau would introduce additional complications. First, the Bureau has a policy of not allowing direct telephone access to its mainframe computers, other than through "dedicated" lines. Even with encryption techniques, use of passwords, and operator intervention, we have concerns about the public perception that computer hackers could get into the live data. A second problem is that if public use files are also created, the test file could potentially be matched to the public use file revealing additional information (data suppressed or modified on the public use file) for those cases on the test file. Also, for unique cases that fall into the test file, removing the geography may not be sufficient to protect the identity of the respondent. Finally, we must be concerned with the possibility that although the individual tabulations are "safe", various combinations, taken together, may reveal unique characteristics about a respondent.

The Census Bureau has recently initiated a Data Resource Center (DRC) for the SIPP which will serve as a testing ground for this approach to disseminating microdata. The DRC was created about two years ago to serve researchers who cannot obtain the data necessary for their analyses from available Census Bureau data products. "It has been designed to serve as both a technical and an administrative link between non-Census Bureau researchers and the data contained on internal Bureau files, especially those files produced from the SIPP data set. Further, it has been charged to coordinate and produce special demographic, social, and economic data sets, tabulations, and analyses for non-Census Bureau researchers and analysts from these files." (Cavanaugh, 1987) Although the Data Resource Center has an ultimate goal of developing useful, and anonymous, test files, so far its primary use has been to provide research files from the SIPP wave data sets. (These research files have been modified to protect confidentiality but have not yet been made public-use because they require further research or evaluation.) Nevertheless, some work has been done on the development of anonymous test files that would be representative of the entire sample. Although much work is required before a Luxembourg-type program is in place at the Census Bureau, the DRC is working with interested analysts to help make it a reality, (Herriot, et.al., 1988).

ADMINISTRATIVE SOLUTIONS

Public use solutions, such as these, will provide benefits for the greatest number of users. They will not satisfy all users and, in particular, may not be the answer for statistical projects funded by other Federal agencies, including our survey sponsors. Many studies requiring the development of models, reprocessing of the raw data, or enhancement with various administrative data sources cannot be done using public use files or summary statistics. The nature of these studies requires use of information that may never be made public use.

Aside from the SIPP, nearly all of the Census Bureau's household surveys are fully or partially sponsored and funded by other Government agencies. The Census Bureau collects and processes the data under a reimbursable agreement and delivers a data product to sponsors (tabulations and/or public use microdata files). Under Title 13, survey sponsors are treated just like other non-Census Bureau employees and are not entitled to see individual records from the surveys they fund. This has presented problems for some of our sponsors--who in fact are primary survey users--and makes it more difficult to fulfill our mission to provide statistical information to a wide variety of users.

Non-Title 13 Surveys

Before 1976, Title 13 did not specifically authorize the Census Bureau to conduct surveys for other Federal agencies. Such work, however, was authorized by the Economy Act (Title 31) that allows one agency to perform work for another agency, or by Title 15, which authorizes the Secretary of Commerce (of which the Census Bureau is a component) to conduct special studies for other organizations. When conducting surveys under these titles, we cited the other agency's authority to collect the data but maintained that the data collected in this manner must be kept confidential when the sample from which the survey was drawn was developed under Title 13 (for example, addresses obtained in the decennial census). On the other hand, if the sample was drawn from lists provided by the sponsor or involved canvassing certain geographic areas (area sampling), the confidentiality, if any, was assumed to extend from the sponsor's authority and not from Title 13. Therefore, respondents were notified that we were acting as a collection agent and that the individual information would be turned over to the sponsor who would protect its confidentiality to the extent permitted by law. When Title 13 was amended in 1976 to give the Census Bureau explicit authority to conduct surveys for other agencies we began to use our own authority and apply the Title 13 confidentiality provisions to all reimbursable surveys conducted under that authority.

With the increasing demand from current and prospective sponsors for identifiable data for use in conducting follow-up surveys or in merging a respondent's individual information with administrative record data, the Bureau established a policy in 1987 to conduct reimbursable surveys under Title 15, rather than Title 13, when the following conditions were met:

1. The sponsor has the legal authority to collect the information and to contract with the Census Bureau for the work.

2. The sample is not derived from Census Bureau records which are protected by Title 13.

3. The purposes, content, methods, or other aspects of the survey are not deemed objectionable by the Census Bureau.

4. The sponsoring agency will: sign an agreement binding the sponsor and its contractors and grantees to use the data only for statistical purposes; notify the respondents of the conditions under which the information is being provided; collect and maintain the data in accordance with applicable Federal laws; and prohibit redisclosure in identifiable form.

We have approximately 12 active surveys conducted under the sponsoring agency's data collection authority. The samples for the majority of these surveys were selected from administrative lists provided by the sponsor. However, we are doing the Health Interview Survey for the National Center for Health Statistics (NCHS) using area sampling and a Point of Purchase Survey (CPP) feasibility test for BLS using random digit dialing. Although we anticipate that we will continue to get requests to conduct surveys outside of Title 13, some sponsors will prefer to use census lists to select their samples because alternative frames are not available or too costly.

Use of Special Sworn Employees

As previously mentioned, the Census Bureau has the authority to use temporary staff to perform work authorized by Title 13. This includes employees of other Government agencies and private organizations. The Census Bureau, at its discretion, can appoint an individual as a Special Sworn Employee (SSE) when: 1) that individual is employed by an agency or organization for which we have a contract to provide services or are engaged in a joint project and the person has expertise or specialized knowledge that can contribute to the accomplishment of our projects or activities; 2) the individual is employed by an agency or organization performing a service for the Census Bureau under contract or provides information to the Census Bureau for statistical purposes; or 3) when Federal law requires an individual to audit, inspect, or investigate Census Bureau activities. As an example, we have sworn in employees of the Social Security Administration (SSA) to obtain information from SSA administrative files about respondents to the SIPP for a matching project that we are jointly undertaking. Also, during each Census of Agriculture, we swear in employees of the Department of Agriculture's National Agricultural Statistics Service to review county level summary data. These experts look for abnormalities in the data, based on their local knowledge.

In 1977, the Census Bureau instituted the ASA/NSF/Census Research Program, jointly funded by the Census Bureau and the National Science Foundation (NSF) and administered jointly by the Census Bureau and the American Statistical Association (ASA). Broadly, the purpose of this program is to promote methodological and substantive research involving Census Bureau databases; to provide hands-on experience for graduate students in the fields of statistics, economics, demography and related areas; and to help bridge the gap between academic and government social

science. The ASA Fellowship Program, as it is commonly called, has been instrumental in bringing improvements in Census Bureau operations--primarily by providing increased communication between Bureau staff and the users of our data. Between 1977 and 1987, 32 Fellows and 25 Associates have participated in the program.

The ASA Fellowship Program has fifteen specific goals to bridge the gap between government and academic social science (Table 1). An Evaluation Conference held in June 1986 found that "the program has been highly successful when assessed in terms of its objectives." Regarding Goal 1 (to provide academic scholars with the unique opportunity to have "hands on" access to Census data), Fellows and Associates have used data sets unavailable to researchers outside the Census Bureau for reasons of confidentiality. For example, several Fellows have used microdata from the SIPP and from the Longitudinal Establishment Data (LED) file. Some participants have used data sets constructed from Census Bureau data and data from other agencies. (ASA Grant Proposal, 1987). Through the various research activities conducted with these data sets many of the other goals of the program have been achieved.

In 1986, we instituted the Interagency Research Fellowship Program which was modeled after the ASA Fellowship Program. This new program, however, was designed to support projects funded by other Federal agencies. We believed that a larger program would expand on the successes already achieved; provide more visibility for the program; stimulate intellectual discussion between Census employees and Government researchers; and open up avenues for new approaches to our problems and procedures. As stated in the proposal for the Interagency Research Fellowship Program, it is intended to:

- foster and stimulate increased use of census data bases for methodological and substantive research which would benefit from access to individually identifiable data;
- provide a research environment emphasizing collaborative interests of the Census Bureau and the social sciences research community; and
- stimulate the exchange of substantive and methodological information between Census Bureau personnel and the academic communities.

To be eligible for this fellowship, a qualified person must have a project acceptable to the Census Bureau. In addition, the project must be funded by another Federal agency, state or local government, or an appropriate research funding source. The project must be accepted as having statistical merit, direct relevance to the Census Bureau mission, and be sponsored by a

component of the Bureau. Finally, the project must be approved by the Director of the Census Bureau who will make his determination based on the merits of the research as well as the long-run benefits and costs that the project may have on other Census Bureau programs. Appointments as Research Fellows are for a period of one year, with continuations of up to three years possible. As with the ASA fellowships, Research Fellows must commit to an extensive period of work at the Census Bureau facilities in Suitland, Maryland.

In the initial application of the Interagency Research Fellowship Program, we have brought in a full time employee of the Economic Research Service (ERS) of the U.S. Department of Agriculture to work with a SIPP file containing non-metro designations. As described previously, this detailed geography was needed to assess the economic well-being of non-metro residents--a study fully supported by the Census Bureau's statistical mission. In addition to the analyst, a programmer for the ERS was assigned for six months at the beginning of the project to create an extract of the specific SIPP file which could be used with SPSS software and also to assist in checking the initial tabulations. The research is proceeding quite well and we expect several reports will be published. Also, there has been a healthy exchange of ideas between the ERS researcher and staff in our Population Division which is supporting the work. Administratively, cost accounting has worked fairly well with a special account set up to draw from the $30,000 allocated for computer expenses. The primary administrative complication resulting from this program is the lack of adequate space in the Division for the Fellows to work. The lack of adequate space may limit the expansion of this program to a great extent, especially until after the 1990 Census.

The requirement of this program that all research with the individual microdata records be done onsite has been a significant limitation to some potential Research Fellows who do not wish to commit so much time away from their homes. Although Title 13 does not require that the data we collect be maintained at a specified facility, it is the Bureau's policy that in order to assure security and maintain the public's confidence, we generally require that the data be used in Suitland. To overcome some of the inconvenience to the Research Fellows and other SSEs, we are experimenting with a procedure to locate restricted data at our regional offices. These offices are located in twelve large cities (Table 2) throughout the United States.

We are experimenting with this program through a Joint Statistical Agreement (JSA) with Harvard University. The purpose of this project is to analyze the results of the Post Enumeration Survey component of the decennial census pretest conducted in Los Angeles, California in 1986. Since the file contains geographic identifications to the block level, a public use product is not possible. In addition, it would be quite inconvenient for the Harvard researchers to come to Suitland to process the data.

As a result, we are providing the individual data from this test to the Harvard researchers, who are SSEs, on a microcomputer located at our Boston regional office which is within commuting distance of Harvard. The computer was brought in by the researcher and the data were loaded from floppy disks. Interactive sessions are restricted to the regional office; however, the tabular output can be analyzed at Harvard. The work is to be done over a period of several months and, upon completion, the computer's hard disk will be scrubbed and the computer will be returned to the University.

In the long run, we would prefer a more centralized approach to this program. We envision dedicating a minicomputer at Suitland for this work and connecting it to each of our regional offices through the secure telephone lines which will support our decennial census activities. Terminals at the regional offices could access specific files for authorized projects. There would be no connections to the Bureau's mainframe computers and the files on the minicomputer would not contain any individual identifiers. Survey data matched with administrative records could also reside on the minicomputer. Staff in Suitland would provide technical support to the Research Fellows by monitoring the interactive sessions.

This regionalized approach will not satisfy those Special Sworn Employees who are great distances from a regional office city; nor will it satisfy some of those located in Washington or near a regional office city who are locked into their own machines due to software requirements or cost factors. However, as in the case of the Harvard JSA, there will be instances where it is preferable given the alternatives.

LEGAL OPTIONS

In addition to our public use and administrative solutions, there are legal options which would extend the obligation to protect confidentiality, and the resulting liability, to the data user. These options involve: (1) creation of statutory penalties for improper use of public use microdata, and (2) legal contracts or license agreements that bind the user of public use microdata to use the data only for prescribed statistical studies.

In support of statutory provisions, Robert Pearson of the Social Science Research Council wrote that: "Acceptable disclosure risks are neither easily nor precisely calculated, but such agencies as the Bureau of the Census and the Internal Revenue Service often require (or interpret the laws that govern the release of such data as requiring) that these levels equal zero. I reveal my prejudices here, if not before, in believing that the extended use of federal statistics per se is not inappropriate; but rather that (1) the value of these data are not fully realized and (2) most current statutes under which the release is governed are

inadequate because they recognize only the obligations of those who collect the information, not the obligations of those who may subsequently use them." (Pearson, 1986).

Similarly, Jelke Bethlehem of the Netherlands Central Bureau of Statistics, in a paper presented at the Census Bureau's Fourth Annual Research Conference, concluded that "...disclosure of micro data sets is possible and often difficult to prevent unless the information in the data set is severely reduced." "Therefore," he wrote, "if micro data are released under the conditions that the data may be used for statistical purposes only and that no matching procedures may be carried out at the individual level, any huge effort to identification and disclosure shows clearly malicious intent. In view of the duty of a statistical office to disseminate statistical information, we think disclosure protection for this kind of malpractice could and should be taken care of by legal arrangements, and not by restrictions on the data to be released." (Bethlehem, et.al. 1988).

There are only a few examples of legal arrangements currently being used by statistical organizations. In West Germany, the Federal Statistical Office assumes that there is a residual risk of disclosure in any release of public use microdata. Consequently, they have a means of releasing microdata to an institution under an agreement requiring that:

- The receiving institution pay the cost of modifying records for disclosure control prior to release;
- The receiving institution not try to reidentify records;
- Data may not be transferred to third parties; and
- Violation of the conditions of the release will result in a fine and exclusion from future access to microdata.

Recently, two laws (the Federal Law of Statistics of January 1987 and the Census Law of November, 1985) have had a significant impact on the release of microdata in West Germany. In the Census Law, Articles 17 and 18 specifically prohibit the reidentification of respondents from census data:

> Article 17
>
> (1) The characteristics, including the block side (Article 15, para. 4, sentence 3), recorded on the basis of this law will be used only for statistical purposes.
>
> (2) It is prohibited to match characteristics pursuant to para. 1, or to combine such characteristics with data from other statis tical surveys, for establishing a reference to individual persons for other than statis-

tical purposes of this law.

Article 18

Whosoever, contrary to Article 17, para 2, brings together characteristics or data after the characteristics according to Article 17, para 1 have been transferred to data media intended for further computer processing, will be liable to a term of imprisonment not exceeding one year or to a fine.

In the United States, two research organizations, Ohio State University's Center for Human Resource Research and the University of Michigan's Institute for Social Research (ISR) are using or are planning to use contracts as a means of releasing more detailed microdata files. Ohio State University releases a public use file from the National Longitudinal Surveys Youth Cohort (NLSY) conducted by the NORC through funds provided by the Department of Labor. In addition to the public use file, a separate "geocode data tape" containing county codes, college identifiers, some administrative data, and limited information from the County and City Data Book are sold to institutions under a license agreement.

The OSU license agreement requires that: 1) results of the research be published only in summary and statistical form such that no individuals can be identified; 2) files will only be used for specified statistical research and will not be released to unauthorized persons; 3) no attempt will be made to identify an individual on the file; 4) the tape recipient may not hold OSU liable for claims resulting from release of the file; 5) the tapes are destroyed when the work is completed; and 6) the recipient agrees to all protections required by the Privacy Act of 1974. This type of agreement has been used since 1980 and there have been no known breaches of confidentiality or evidence of impropriety. Presumably, if a breach were to occur, the main recourse to OSU is to stop providing the guilty user with these kinds of microdata.

The ISR proposal involves the Panel Study of Income Dynamics (PSID) which ISR conducts with funds provided by the National Science Foundation and others. Currently, PSID public use files show geography to the county level (there are no restrictions on county size). To meet increasing demands for local area data, special research files are being created which identify records by census tract and by ZIP code. ISR plans to release these files to researchers whose institutions co-sign an agreement patterned after the Ohio State license agreement. The recipient institution would be required to provide a detailed proposal as to how they plan to protect the data while it is in their possession. If ISR agrees that the measures are appropriate, the researcher must post a $1000 security deposit before the files would be released. Upon completion of the work, the recipient

attests that all files or derivatives have been destroyed and signs a statement that no known breaches of confidentiality have occurred. The $1000 deposit would then be returned.

I know of no examples of statutory penalties or legally binding contracts regarding the release of microdata currently in use by U. S. Federal Statistical Agencies. The National Center for Health Statistics (NCHS) does, however, require purchasers of public use microdata tapes to sign a statement in which they agree to abide by the NCHS legislation which states that "the data may be used only for the purpose for which they were obtained, i.e., for statistical purposes." (Mugge, 1983) This signed statement is in addition to established disclosure protection measures which are similar to those employed by the Census Bureau. Although not a means to provide greater access, the statement does help to sensitize the user to NCHS' concern for the confidentiality promised to the respondent.

A legislative approach that could expand access to Census Bureau microdata involves creating a new type of appointment, similar to Research Fellowships, that would provide access to microdata only for general statistical research. Persons, so appointed, would not be Census Bureau employees and would not be restricted to doing research specifically tied to Census Bureau work. This would open the Research Fellowship Program to additional researchers and would remove the time limitations associated with temporary employees (SSEs).

Currently, contractual and legislative options such as user liability and research appointments are not available to the Census Bureau. Title 13 does not give us the option of sharing liability with microdata users or providing access to identifiable records by non-Census Bureau employees. The Census Bureau will look at legislative changes as a means of supplementing or replacing our public use and administrative programs. If such solutions are deemed appropriate, we would need to carefully evaluate how our respondents would react to sharing our responsibility for protecting their data with others before we recommend any modification to Title 13. In addition, we would have to consider the sensitivity of the information on the file and the consequences of a possible breach on our ability to gain the voluntary cooperation of our respondents in the future.

CONCLUSIONS

This year marks 25 years of producing public use microdata files. When we originally conceived the idea, we thought that most users would want to receive the information on computer punch cards, (Zeisset, 1988). Things have changed a lot in these 25 years--to the point where over one-half billion bytes of information can now be stored on a single CD-ROM disk. Now, many private researchers and the staffs of nearly all Government agencies have

the ability to process large databases and to apply sophisticated analytical and modeling techniques. The potential social and economic benefits resulting from this research are enormous.

On the other hand, public concerns for protecting individual privacy and confidentiality have been heightened by the vast databases maintained by Government agencies and the trend toward matching files across agencies. These practices, along with the ease with which the data can be handled and analyzed, may cause survey respondents concern that the Government may use their responses, that were provided voluntarily, against them in some way. Businesses, on the other hand, are concerned that competitors will take advantage of information they may glean about their financial situation or marketing strategies. These concerns, if substantiated by a misuse of statistical data, could reduce participation in our censuses and surveys. Although this result may not affect the immediate short-term goals of any individual researcher, it would certainly be a long-term tragedy for the entire statistical community and should provide sufficient incentive for researchers not to abuse the trust respondents placed in the Census Bureau when they provided the information. However, even an innocent misuse or carelessness may be all that is required to markedly reduce public participation in our programs. Also, public use products are not restricted to bonafide researchers and others may not be so motivated.

In this environment, the Census Bureau, as a service organization, must continue to provide the best possible service to our users--especially Federal users who depend on our data to make policy decisions that affect the quality of life for millions of Americans and, at the same time, are responsible for allocating billions of taxpayer dollars. In addition, we must continue to examine and evaluate the potential risks of identifying survey and census respondents from public use microdata and we must establish criteria for acceptable levels of risk. Where public use microdata are not possible given this risk, we will consider alternative products and administrative arrangements that satisfy our user's statistical requirements. Finally, and perhaps most importantly, we **MUST** ensure our respondents that the data they provide the Census Bureau for statistical purposes will **NOT** be used to make determinations about them as individuals.

Acknowledgement

I wish to thank Sherry Courtland who provided guidance throughout the development of this paper and particularly with its organization. I am also indebted to Brian Greenberg and Sang Nguyen for their thorough review and helpful recommendations, especially in the areas of disclosure risk and reduction and public use alternatives to microdata.

REFERENCES

ASA Grant Proposal submitted to the NSF, March 1987; "On-site Research to Improve Government Generated Social Science Data Base;" Proposal SES-8713643.

Bethlehem, J. and W. Keller and J. Pannekoek (1988), "Disclosure Control of Micro Data;" paper presented at the Census Bureau's Fourth Annual Research Conference, Arlington, Virginia; March 1988.

Butz, W.P. (1985), "Data Confidentiality and Public Perceptions: The Case of the European Censuses," 1985 Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 90-97.

Cavanaugh, F.J. (1987), "SIPP As An Initiator of a Data Resource Center At the Census Bureau." 1987 Proceedings of the Section on Statistical Computing, American Statistical Association, pp. 149-154.

Dalenius, T. (1988), Controlling Invasion of Privacy in Surveys, Statistics Sweden, Garnisonstryckeriet Stockholm, 1988.

Dalenius, T. (1988), "The Debate on Privacy and Surveys in Sweden," Chance: New Directions for Statistics and Computing, Vol. 1, No. 2; Spring 1988; pp. 43-47.

Duncan, G. and D. Lambert (1987), "The Risk of Disclosure for Microdata," Proceedings of the Third Annual Research Conference, U.S. Bureau of the Census, April 1987.

Govoni, J. and P. J. Waite, (1985), "Development of Public Use File for Manufacturing," 1985 Proceedings of the Section on Business and Economic Statistics, American Statistical Association, pp. 300-302.

Greenberg, B. (1988) "Disclosure Avoidance Research at the Census Bureau," paper presented at the Census Bureau's Joint Advisory Committee Meetings, Arlington, Virginia; April 1988.

Herriot, R. and C. Bowie, D. Kasprzyk, S. Haber (1988), "Enhanced Demographic Data Sets with Employer Characteristics," paper presented at NBER Conference on Research in Income and Wealth, Washington D. C. May 12-14, 1988.

Kim, J. (1986), "A Method for Limiting Disclosure in Microdata Based on Random Noise and Transformation," 1986 Proceedings of the Survey Method Research Section, American Statistical Association, pp. 370-374.

McGuckin, R.H. and S. Nguyen, (1988), Use of Surrogate Files to Conduct Economic Studies with Longitudinal Microdata"; paper presented at the Census Bureau's Fourth Annual Research Conference, Arlington, Va.; March 1988

Monahan, J. (1986), "Development of Microdata Public Use Data File from the Longitudinal Establishment Data File." Internal Memorandum, Center for Economic Studies, U. S. Bureau of the Census.

Mugge, R.H. (1983), "Issues In Protecting Confidentiality In National Health Statistics," 1983 Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 592-594.

Paass, G. (1985), "Disclosure Risk and Disclosure Avoidance for Microdata," working paper, Institute for Applied Information Technique, Gesellschaft fur Matematik und Datenverarbeitung, Sankt Augustin bei Bonn, Federal Republic of Germany.

Pearson, R.W. (1986), "The Confidentiality and Extended Use of Federal Statistics," paper presented at the American Statistical Association Meetings, Chicago, Illinois, August 1986.

Rainwater, L. and Smeeding, T. (1988), "The Luxembourg Income Study: the Use of International Telecommunications in Comparative Social Research," The Annals of the American Academy of Political and Social Science, January 1988.

Spruill, N.L. (1984), "Protecting Confidentiality of Business Microdata by Masking," The Public Research Institute.

Wolf, M. (1988), "Microaggregation and Disclosure Avoidance for Economic Establishment Data," paper for presentation at the Joint Session of the Business and Economics Statistics Section and Section on Survey Methods Research, American Statistical Association Meetings, August 1988.

Zeisset, Paul (1988), Interview at Bureau of the Census, Suitland, Md.; May 11, 1988.

Table 1

OBJECTIVES OF THE ASA/NSF/CENSUS RESEARCH PROGRAM

1. To provide academic scholars with the unique opportunity to have "hands on" access to Census data.

2. To provide increased opportunity for accomplished social scientists to work on important problems in a non-academic environment, where production and research needs are often different and can conflict.

3. To stimulate methodological and substantive research in academia on the problems of collecting and analyzing data that provide the basic information for making decisions that can have broad impacts on society.

4. To increase exposure of Census Bureau social scientists to outside expertise, and hence to broaden their perspectives regarding the ultimate analytic uses of the data they produce.

5. To bring about an improvement of the quality of the data collected and disseminated by the Census Bureau.

6. To further specific scientific advances in methodological and substantive areas related to the data collection activities of the Census Bureau.

7. To provide an opportunity for graduate training and doctoral dissertation research using the problems of governmental data collection agencies.

8. To develop a resource group of personnel for future recruitment of statisticians and social scientists to help fill governmental research needs.

9. To provide a large variety of usable real data, as well as computer software programs for their analyses, for teaching and research at academic institutions.

10. To conduct seminars and conferences jointly sponsored by a group of agencies and academic institutions.

11. To increase the interaction and collaborative research and education among agencies and between agencies and academic institutions.

Table 1 (Cont.)

12. To improve the quality of the statistical analysis of Census Bureau data.

13. To suggest important new analyses of existing data that can and should be done.

14. To generate a positive impact on curriculum development at academic institutions.

15. To develop a cadre of people experienced in problems of data methodology and data use who will submit high-quality proposals to NSF to pursue basic and applied research based upon novel ideas and approaches.

Table 2

CENSUS BUREAU REGIONAL OFFICES

Boston, Massachusetts

New York, New York

Philadelphia, Pennsylvania

Detroit, Michigan

Chicago, Illinois

Kansas City, Kansas

Seattle, Washington

Charlotte, North Carolina

Atlanta, Georgia

Dallas, Texas

Denver, Colorado

Los Angeles, California