DRAFT:
NOT FOR CITATION
OR QUOTATION

THE EFFECT OF ATTRITION ON INCOME AND POVERTY ESTIMATES FROM THE SURVEY OF INCOME AND PROGRAM PARTICIPATION (SIPP)

by

Enrique Lamas, Jan Tin, and Judith Eargle
U.S. Bureau of the Census

May 4, 1994

Enrique Lamas, Jan Tin, and Judith Eargle are with the Housing and Household Economic Statistics Division, U.S. Bureau of the Census. The authors want to thank Jeffrey Zabel (Tufts University and Census Bureau) for many fruitful discussions about the issues discussed in this paper and for his help with the estimation techniques and programs. This paper was presented at the Conference on Attrition in Longitudinal Surveys, February 24-25, 1994, Washington DC. It reports the general results of research undertaken by the Census Bureau staff. The views expressed are attributable to the authors and do not necessarily reflect those of the Census Bureau.

THE EFFECT OF ATTRITION ON INCOME AND POVERTY ESTIMATES FROM THE SURVEY OF INCOME AND PROGRAM PARTICIPATION (SIPP)

by

Enrique Lamas, Jan Tin, and Judith Eargle
U.S. Bureau of the Census

The Survey of Income and Program Participation (SIPP) provides federal policy makers and researchers with detailed income and poverty data on the Nation's households. SIPP is designed as a longitudinal survey where individuals are interviewed at relatively frequent intervals (every four months) for a period of two and two-thirds years. The longitudinal nature of the SIPP provides the opportunity to examine movements along the income distribution and factors associated with exits from poverty or with the duration of spells in poverty. For example, data on exits from poverty from the SIPP showed that, overall, 21 percent of persons below poverty in 1990 were not poor in 1991. However, the exit rate was higher (42 percent) for persons who worked year-round, full-time [U.S. Bureau of the Census (1993)]. The SIPP also provides the opportunity to examine the effects of changes in family composition on income and poverty estimates and supplementary data provide the ability to examine related issues such as asset spend-down of persons in poverty [Ruggles and Williams (1989)].

Attrition is, however, known to be a problem in longitudinal surveys. A few papers have examined the problem of attrition in SIPP [Short and McArthur (1986); Lepkowski, Miller, and Luis (1993)]. The extant work, however, has primarily used the initial panels from SIPP (1984-1986) and has not focussed on income and poverty estimates. The purpose of this paper is to examine the effects of attrition on income and poverty estimates derived from SIPP and to present some possible methodologies for improving the estimates.

In the first section of this paper, we focus the study on estimates of the number and percent of persons in poverty. We compare longitudinally-weighted poverty estimates from SIPP to the official poverty estimates from the Current Population Survey (CPS). The poverty estimates from SIPP are substantially lower than estimates from the CPS. We explore several methodological differences between the surveys to determine how much each of these factors contribute to the differences in the poverty estimates. These include differences in the treatment of household composition, differences in self-employment income, and differences in the collection of program participation information. We find that the differences in survey methodologies do not account for all the differences in the poverty estimates. Therefore, we examine the effect of attrition on income and poverty estimates as a possible source of the differences.

In the second section of the paper, we examine the demographic and economic characteristics of attritors. In addition, in a multivariate framework we examine the relationship between events, such as marital status or employment status changes, and the likelihood of attrition. We also examine several patterns of nonresponse including those respondents who missed some interviews and were later interviewed, and those who attrited fully from the survey.

In the third section, we examine the effect of attrition on income and poverty correlates using several models of income and poverty that take attrition into account. Equations are estimated for total income, labor income, non-labor income, means-tested income, and poverty status in order to test whether attrition has an effect on income and poverty as well as to examine the effect on the behavioral parameters.

In the final section, we examine the magnitude of potential attrition bias on poverty estimates using simulations. We impute missing information for attritors and calculate poverty estimates for the complete panel. To obtain an estimate of the potential attrition bias, we compare poverty estimates for the full panel using simulations for attritors to that of panel members with complete information.

Poverty Measures

The official income and poverty estimates from the Federal Government are issued by the Census Bureau based on data from the annual March Demographic Supplement of the CPS. These estimates fix family composition as of the survey date (March) and ask about income received during the previous calendar year. Poverty status is determined by comparing annual family income to the appropriate annual poverty thresholds based on the size of the family and number of related children under 18.[1] In SIPP, income information and family composition are collected on a monthly basis. Monthly poverty thresholds are constructed by prorating the official annual thresholds to a monthly basis (dividing by 12) and adjusting them for monthly changes in the Consumer Price Index (CPI-U). We construct annual poverty estimates by comparing the sum of monthly family income to the sum of monthly poverty thresholds over the year based on the family characteristics in each month.

Poverty estimates vary considerably between the official CPS estimates and the SIPP estimates.[2] Table 1 presents poverty estimates for

---

[1] For a discussion of the history of poverty estimates from the CPS and current issues related to the measurement of poverty, see Ruggles (1990), and Weinberg and Lamas (1994).

[2] In making these comparisons we recognize that the CPS and SIPP suffer from various problems typical of household surveys, such as underreporting of income, household and item nonresponse, and coverage problems. However, the comparisons between the surveys are useful because: (1) there is no other source of independent administrative sources for benchmarks of poverty estimates; (2) both CPS and SIPP are adjusted to the same population estimates; and (3) the CPS is the

1990 from the CPS and SIPP. The SIPP estimate of the percent of persons below the poverty level in 1990 is 10.5 percent or 3.0 percentage points lower than the CPS poverty rate (13.5 percent). The SIPP estimates are also considerably lower than CPS estimates when we examine characteristics such as age, race, and marital status.

There are several differences between the two surveys that contribute to the difference between the SIPP and CPS poverty estimates. First, the CPS and SIPP poverty measures are conceptually different in the treatment of household composition. As described above, in the CPS the household composition is fixed as of March and their income over the previous year is used even though the household composition may have changed during the year. The SIPP allows for income and family composition to vary on a monthly basis. In order to examine the effect of this difference, we used the SIPP monthly data to replicate the CPS family definition. We fixed household composition as of March 1991 and calculated poverty status by comparing the income received in the previous year by those persons present in March to a poverty threshold based on the family composition as of March.[3] Even when the SIPP data are used to replicate the CPS cross-sectional measure of poverty (SIPP-CS in Table 1), this revision explains only about one-sixth of the difference between the SIPP longitudinal estimate and the CPS estimate.[4] We found a similar result when we examined poverty estimates by characteristics of persons; the SIPP cross-sectional estimates explained only a small proportion of the difference between the SIPP and CPS poverty estimates by age, race, sex, and educational attainment. Therefore, the conceptual difference in family definition between the SIPP and the CPS does not completely account for the difference in poverty estimates.

A second methodological difference between the surveys is the treatment of self-employment income. The SIPP uses a draw or salary concept from businesses to collect self-employment income and does not allow reporting of losses. The CPS definition uses a net income approach which allows for losses. This difference in the treatment of self-employment income would result in higher poverty estimates from CPS. In order to estimate the effect of this difference we estimated poverty rates from CPS that excluded cases with negative self-employment income. We found that exclusion of losses from CPS would have only reduced the poverty rate from 13.5 to 13.4 percent in 1990 and accounted for only a small part of the difference.

A third important difference between the surveys is that the SIPP has improved reporting of program income. The SIPP design, including its subannual reference period, is better suited for reporting program participation data than the CPS design. Studies

---

official source of Federal statistics on poverty.

[3]In order to replicate the CPS weighting, we used cross section weights for March 1991 for these SIPP estimates.

[4]Coder, et. al. (1987) found similar results when they compared SIPP and CPS estimates covering a 12-month period from late 1983 to 1984 using the first longitudinal data file from SIPP produced from the 1984 panel.

comparing the SIPP and CPS using the 1984 panel found that SIPP reported more recipients of cash means-tested income, reported more aggregate income, and was closer to independent benchmarks than CPS [Vaughan (1989)]. However, work on the 1990 panel found that SIPP identified more program recipients, but it did not report more aggregate income than the CPS for some programs [U.S. Bureau of the Census (1994)]. To the extent that SIPP has better reporting of means-tested cash income, it would tend to reduce poverty estimates in SIPP.

Finally, the SIPP longitudinal estimates presented above are based on persons for whom a complete set of observations were obtained. The SIPP estimate shows the 1990 poverty status for the cohort of persons who lived in U.S households in January 1990. The SIPP weighting procedures do not give positive weights to persons who missed one or more interviews. This includes persons who missed one interview as well as persons who attrited and did not return to the survey. If the persons with missing data or those that moved in with original sample persons were more likely to be in poverty, then the poverty estimates would be biased. Hill (1993) examined the effects of weighting on poverty estimates and found that cross-sectional estimates for a month were significantly higher than estimates using longitudinal weighting for original panel members with complete information over the panel. Hill suggests that persons who joined the sample or persons with missing interviews were more likely to be poor, that is, that there is nonrandom attrition with respect to poverty. The next section of this paper examines the characteristics of persons who attrit from the panel and examines the effect of attrition on poverty estimates.

Characteristics of Attritors

The SIPP panel design interviews households every four months for a 32-month period. The 1990 panel collects monthly data on approximately 58,300 persons in the United States based on interviews conducted from February 1990 to September 1992. The civilian noninstitutional population of the United States and members of the Armed Forces living off post or with their families on post are covered by the SIPP. The primary focus of SIPP are persons 15 years old and over who were interviewed in the first wave of the panel. These "original sample persons" are followed over the life of the panel. If original sample persons move during the life of the panel, they are followed to the new address and all persons residing with them are interviewed. Persons added to the sample because they live with original sample persons are followed until they no longer reside with original sample persons [U.S. Census Bureau (1991)]. One source of nonsampling error is attrition of original sample persons. If panel attrition is a non-random event and related to certain characteristics or events, it may result in biased estimates of those characteristics or events of interest [Hausman and Wise (1979)].

For the purposes of this paper, we define attrition to be original sample persons missing one or more interviews whether or not they return to the sample. Excluded from this group are cases that left the universe of the sample (primarily those who died or became institutionalized during the life of the panel) and therefore cannot be used to estimate poverty status because they do not have complete data for the period.

As with most household panel surveys, most of the sample loss in SIPP occurs in the beginning of the panel. About half of the sample loss occurs in the first three waves of interviews and after the fourth wave most of the sample loss is due to movers [Jabine (1990)]. Household nonresponse due to refusal to participate or due to not being able to conduct an interview after repeated contacts was 7.1 percent of eligible households in the first wave of the 1990 Panel of SIPP. The sample loss was approximately 16 percent by the fourth interview and increased slightly to 20 percent by the end of the panel.

To examine the effect of attrition from the panel, Table 2 presents the characteristics of non-attritors (those who completed all interviews), attritors (original sample persons who missed one or more interviews) and those that left the universe (died, became institutionalized, joined the Armed Forces, or went overseas). Approximately 75 percent of all original sample persons were interviewed in all waves. The results, however, suggest that there are some systematic differences between attritors and non-attritors in demographic and economic characteristics. Goodness-of-fit Chi Square tests of the distributions show that there are differences between attritors in their distributions of age, race, sex, educational attainment, and other characteristics. The distributions show that attritors are more likely to be young adults, males, in minority groups, never married, of lower educational attainment, and in poverty than non-attritors. For example, a greater proportion of attritors than non-attritors were young adults (34.1 percent were 18 to 34 years of age compared to 25.2 percent of non-attritors) and a smaller proportion of attritors were children (21.5 and 25.3 percent, respectively, were less than 15 years) or elderly (7.2 versus 11.5 percent were 65 years and older). In addition, 50.3 percent of attritors were male compared to 46.6 percent of non-attritors; 75.1 percent of attritors were White and 20.7 percent were Black compared to 84.5 percent and 12.3 percent, respectively, for non-attritors; about 49.9 percent were never married compared to 42.5 percent of non-attritors. There were also some differences in poverty status. Poverty status depends on family income as well as needs (poverty thresholds) and there were differences between attritors and non-attritors in the distribution of each. For example, attritors were more concentrated at the lower end of the distribution with 20.9 percent of attritors having monthly family income less than $1,000 compared to 15.7 percent for non-attritors. As a result, 17.1 percent of attritors had income below the poverty level compared to 13.4 percent of non-attritors.

Persons that left the universe were also different than attritors and non-attritors. They were more likely to be elderly, be widowed/divorced/separated, have less than high school education, and to not be in the labor force than either attritors or non-attritors. This is not surprising because the process of leaving the universe (such as through death or institutionalization) is clearly different than the decision to participate or not in the survey. For this reason, we have excluded cases that left the universe from the analysis in the remainder of this paper where we model attrition.

In order to examine in a multivariate context whether these characteristics and other events were correlated with the probability of attrition, we estimated probit regressions of attrition. We used a reduced-form equation of attrition as a function of demographic characteristics (age, race, sex, family and marital status), economic characteristics (education, labor force status, income level), and specific events or changes in these characteristics over the life of the panel which could affect the likelihood to attrit (mover, changes in education, changes in labor force and marital status). Specifically, we estimated a probit equation of

$$a_i = \beta' x_i + \epsilon_i \tag{1}$$

where $a_i = 1$ if person i missed one or more interviews,
$a_i = 0$ if person completed all interviews,
X_i is a set of independent variables related to attrition and ϵ_i is a random error term.[5] We estimated a model for all original sample cases and then estimated separate models for cases with income below the poverty level at the end of the panel or at the time of attrition from the panel. In this way, we allowed for full interaction effects between poverty status and the variables in the attrition equation.

The results in Table 3 show that for all respondents several variables were correlated with attrition. Attrition was positively related to the years of schooling and to income level. Females and married persons were less likely to attrit, while Blacks, Hispanics, and disabled persons were more likely to attrit. Also some changes in status during the panel had an effect on attrition. Changes in employment (either full- or part-time), changes in marital status, and increases in years of education during the life of the panel reduced attrition. As expected, moving also had a significant and positive effect on the likelihood to attrit because of the difficulties in tracking movers within the interview month. In addition, fewer family ties within the household (that is, being a nonrelative of the reference person) had a positive effect of likelihood to attrit.

---

[5] The standard probit model has the underlying response equation

$$a_i^* = \beta' x_i + \mu_i$$

where a_i^* is th likelihood of attrition. However, we only observe whether a person attrits or not, that is,
$a_i = 1$ if $a_i^* > 0$ (or $\mu_i > -ßX_i$), and
$a_i = 0$ otherwise.
From this model, we get that
$\text{Prob}(a_i = 1) = \text{Prob}(\mu_i > -ßX_i) = 1 - F(-ßX_i)$
where F is the cumulative density function of μ_i and μ_i is distributed $N(0,s^2)$ [Maddala (1983)].

While the attrition equation for all persons showed that the likelihood of attriting was correlated with various socio-economic characteristics, the attrition equation estimated for persons below the poverty level showed fewer factors associated with attrition. Attrition for persons below the poverty level was not related to age or marital status. Attrition was related to race, ethnicity, and years of education; poor Blacks and Hispanics were more likely to attrit than poor Whites or non-Hispanics, and poor persons with less education were more likely to attrit. In addition, attrition was positively related to a change in full-time employment, that is, if a poor person changed employment status they were more likely to attrit.

Effect of Attrition on Income and Poverty Estimates

A. Methodology

While the results above indicate some systematic differences in demographic and economic characteristics between attritors and non-attritors, we need to consider how nonrandom attrition affects specific variables of interest, specifically, income and poverty estimates. In this section, we examine several approaches to test for and estimate the effect of attrition on income and poverty.

In one of the most comprehensive studies on attrition in the Panel Study of Income Dynamics (PSID), Becketti, Gould, Lillard, and Welch (1988) examined the representativeness of the PSID over its first 14 years (from 1968 to 1981). They studied the demographic characteristics of attritors compared to a cross-sectional survey (the CPS) and tested whether attrition had an effect on earnings. Specifically, they considered whether attritors and non-attritors were different in terms of the earnings relationships at the beginning of the panel. As part of their analysis, they estimated earnings equations using earnings data at the start of the panel (1968 data) for those respondents in the initial sample, those who remained in the sample in 1975, and those still in the sample in 1981. To examine whether differences between the initial sample and the sample after 14 years had an effect on earnings coefficients, they tested for differences in the earnings equations. They found that earnings coefficients were similar for stayers and for the respondents in the initial sample and concluded that there was little evidence that attrition was systematically related to earnings.

Hausman and Wise (1979) were one of the first to develop a method to correct for attrition in a model and to provide a test for the effect of nonrandom attrition on model estimates. Their work expanded the selectivity bias approach from the cross-sectional data context [Heckman (1979)] to a multi-period context of panel data. Hausman and Wise used their methodology to analyze the earnings response in the Gary Income Maintenance Experiment before and after the experiment treatment (two periods). Since attrition for some groups was substantial, they developed a methodology to correct for attrition from the experiment. Ridder (1990) generalized their methodology to consider the problem of attrition over multiple waves. Essentially, the Hausman/Wise and Ridder methodologies estimate a structural model of the variable of interest (e.g., earnings) and an attrition equation for the sample. Using the results of the attrition equation, they test for selectivity bias and correct for nonrandom attrition bias in the structural equation. Their methodologies are developed in more detail below.

B. Empirical Results

We begin the analysis with an approach similar to Becketti, Gould, Lillard, and Welch (1988). For income, we estimated models of income at the beginning of the panel separately for original sample persons and for the subset of original sample persons who remained in the sample in month 32. We tested for significant differences between sets of coefficients associated with the determinants of income using F-tests. We used a standard model for the income equation. The dependent variable was the log of total monthly income in the first month of the panel. Independent variables included age and age squared (to capture the life cycle pattern of income), years of education, marital status (binary variable for married couples), race and ethnicity (binary variables for non-White and Hispanic origin), health condition (binary variable for persons with a work disability), and geographical location (region and metropolitan residence). Since we are also interested in examining the effect on poverty, we experimented with using three measures of income including non-labor income and means-tested income as dependent variables.[6] Table 4 presents the results for all persons and Table 5 presents similar estimates for persons below poverty.

For each income equation, the coefficients were significant and had the expected signs. The coefficients for age were consistent with the life-cycle hypothesis (increasing during work years and declining in later years). Education had a positive effect on income as expected, while female, disabled and married persons had lower income than males, non-disabled or other marital statuses. The West and metropolitan areas had higher incomes relative to the South and non-metropolitan areas, respectively. Results were similar for means-tested income, except that disabled persons had higher means-tested income (since some programs such as SSI are designed for disabled persons) and all regions had higher means-tested income than the South.

When we examined the differences in the coefficients for the effect of attrition by comparing the results of stayers to the original sample, we found few significant differences between samples. F-tests for differences in coefficients values were significant only for non-labor income, but they did not indicate any significant differences in the coefficients for initial total, labor, or means-tested income between the original sample persons and those remaining in month 32.[7]

---

[6]Means-tested income are government transfers that have complex eligibility requirements for income and other economic resources. These programs include Federal and State Supplemental Security Income (SSI), Aid to Families with Dependent Children (AFDC), general assistance, and Veterans' pensions.

[7]We also included a series of binary variables in the income equation to estimate the time dimension of attrition. Becketti, et. al. (1988) used this approach in their earnings equation. We did not find any significant effect for attrition variables which was consistent with our previous results.

We estimated a similar set of equations for persons below the poverty level (Table 5) and the coefficients were significant and had the expected signs. Attrition of poverty level cases had an effect on the coefficients of initial total income, nonlabor, or means-tested income. The results from this approach suggest that there were some differences between attritors and non-attritors in terms of income and its determinants at the start of the panel.

We used a similar approach to test for the effect of attrition on poverty status. Since poverty status is a binary variable, we adapted the approach and used a probit estimation of initial poverty status as a function of socio-economic variables. In order to develop a test for differences in the coefficients (similar to the F-test for ordinary least squares (OLS) regressions), we estimated separate probit models for initial poverty status of all respondents, attritors, and non-attritors, and used a likelihood ratio to test for differences in the coefficients.[8] The results are presented in Table 6.

Females, Blacks, Hispanics, and the disabled were more likely to be poor, while married persons, and those with higher education were less likely to be poor. Persons in the Northeast and in metropolitan areas are also less likely to be poor than persons in the South or in non-metropolitan areas. Similar results were found when the model was estimated for attritors and non-attritors. When we tested for differences in coefficients using the likelihood ratio test, we found significant differences between attritors and non-attritors in terms of the characteristics related to poverty status at the start of the panel.[9]

While the approach above examined the representativeness of the remaining sample in terms of their characteristics at the start of the panel, we also examined the effect of attrition by estimating equations of income and poverty using the methods developed by Hausman and Wise (1979) and Ridder (1990). These methods provide a test for nonrandom attrition not only at the start of the panel, but during

---

[8] The likelihood-ratio test is

$$\chi^2(q) = 2(L_1 + L_2 - L)$$

where q = the degree of freedom or the number of restrictions on the parameters,
L = likelihood value for the full sample regression,
L_1= likelihood value for the non-attritor regression, and
L_2= likelihood value for the attritor regression.

[9] We also estimated initial poverty models for Whites, Blacks, and Hispanics separately. This method provided interaction effects between race and each of the independent variables. The results using the likelihood ration tested show differences in the effect of independent variables on initial poverty status between attritors and nonattritors within each race and ethnicity group.

the life of the panel. We used the standard attrition model consisting of an income equation and an attrition equation [Hausman and Wise (1979) and Ridder (1990)] where

$$y_{it} = \beta' x_{it} + \epsilon_{it} \tag{2}$$

and

$$a^*_{it} = \alpha_0' w_{it} + \alpha_1 y_{it} + \mu_{it} \tag{3}$$

where i indexes individuals and t indexes time periods. Equation (2) is the income equation where y is log of real income, X is a set of variables determining income, and ß is a set of parameters to be estimated. The residual ϵ is assumed to be normally distributed with zero mean and constant variance.[10] In equation (3), a_{it}^* is the tendency to attrit for individual i at time t and is assumed to be a function of income and a set of exogenous variables (w_i) with parameters, α_o. The effect of income on attrition is captured by α_1. The error term, μ_{it}, is assumed to have zero mean and constant variance. The tendency to attrit, a^*_{it}, is a latent variable which cannot be observed. Nonetheless, the actual attrition, a_{it}, is observable and can be used as a proxy for the tendency to attrit. We assume

$$a_{it} = 1 \quad if \quad \alpha_0' w_{it} + \alpha_1 y_{it} > -\mu_{it} \tag{4}$$

and

$$a_{it} = 0 \quad if \quad \alpha_0' w_{it} + \alpha_1 y_{it} \leq -\mu_{it} \tag{5}$$

Equations (2) and (3) are simultaneously determined. A change in the explanatory variables in the income equation indirectly influences the likelihood of attrition. In addition, if the error terms ϵ_{it} and μ_{it} are correlated, ordinary least squares estimates of the coefficients of income are biased [e.g. Hausman and Wise (1979)]. Heckman (1976, 1979) showed that consistent estimates of the income coefficients can be obtained by using a two-step estimation procedure. First, substituting equation (2) into equation (3) we get the reduced-form attrition equation

---

[10] The error therm also includes an individual effect and a time effect [Hausman and Wise (1979)]. Further decomposition does not alter the results presented in this study.

$$a_{it} = \phi_0' z_{it} + \alpha_1 \epsilon_{it} + \mu_{it} \qquad (6)$$

where ϕ_o' is a vector of reduced form parameters and z_{it} is a set of exogenous variables for the ith individual at time t. The coefficient estimates of equation (6) can be obtained by applying a maximum likelihood probit procedure. This procedure also yields an estimate of the ratio between the standard normal probability density function and the cumulative density function which can be used in the income equation to correct the bias

$$E(\epsilon \mid \phi_o' z_{it} > -\alpha_1 \epsilon - \mu) = \frac{\delta_{\epsilon, \alpha_1 \epsilon + u}}{\delta_{\epsilon}} \frac{f(\phi_o' z_{it} / \delta_{\epsilon})}{F(\phi_0' z_{it} / \delta_{\epsilon})} = \delta \lambda_{it} \qquad (7)$$

where f and F are the probability and cumulative distribution functions, respectively. The ratio of these two functions is represented by lambda. The δ's are the standard errors or covariances of the error terms.

In the second step, the estimate of this ratio is inserted as an independent variable into the income equation (2) as a bias- correcting variable. Applying the OLS method then yields consistent estimates of the coefficients. The final form of the income equation can be rewritten as

$$y_{it} = \beta' x_{it} + \delta \hat{\lambda}_{it} + \epsilon_{it} \qquad (8)$$

where lambda hat is the estimate of the ratio between the standard normal probability density function and the cumulative density function, and δ is the coefficient of lambda.

In past studies on attrition, researchers examined the effect on earnings [e.g., Hausman and Wise (1979)]. Since the main focus of this paper is on persons below the poverty level, all cash income is taken into account.[11] As in the earlier part of this study, we examined total, labor, non-labor, and means-tested income and estimated separate regressions for all original sample persons and for persons

[11] Hausman and Wise assume that non-labor income is exogenously determined. This assumption is not realistic in our study since the share of non-labor income in the total income of many poor persons is likely to be large relative to that of labor income.

below the poverty level.[12] We used the model of income similar to one used above which had log of income as a function of demographic and economic characteristics. We also used the model of attrition similar to the one in the previous section where attrition was a function of demographic and economic characteristics. However, the independent variables of the attrition equation are lagged one period. [13]

As shown in Table 7, there also was no evidence that attrition had a significant effect on income (the coefficient of lambda was not significant for any income type.) Most coefficients in the income equations were statistically significant and had the expected signs. For example, the coefficient of age was positive and the coefficient of age squared was negative, consistent with the life-cycle hypothesis. Education was positively related to total income, while females, Blacks, and Hispanics had lower incomes than males, Whites, and non-Hispanics, respectively. Persons in the West and in metropolitan areas had higher income than persons in the South and non-metropolitan areas. The results for the labor,[14] non-labor, and means-tested income were similar and there was no evidence that attrition had a significant effect on income.

We also examined the effect of attrition on the income of persons below the poverty level. As shown in Table 8, we found some evidence of nonrandom effects of attrition on total and non-labor income (the coefficient of lambda was significant), but there was no significant effect on labor income or means-tested income. There were also some differences in the results for specific variables from those of all persons. Age and race were not statistically significant for any of the income types. Education had a negative effect on total and non-labor income, but no effect on other incomes.

We also used a selectivity bias approach to examine the effect of attrition on poverty status. Since poverty status is a binary variable, the estimation technique for the structural equation was modified in order to obtain consistent coefficient estimates. The structural equations for poverty status and the probability of attrition are

$$PV_{it}^{*} = \beta' X_{it} + e_{it} \tag{9}$$

and

---

[12]Since we examined income over the 32-month period, we used real income standardized using the Consumer Price Index (CPI-U).

[13]When attrition occurs, we do not obtain an observation for that period.

[14]In this section we were interested in testing whether attrition had an effect on the behavioral coefficients of labor income. We did not examine the effect of selectively bias on labor force participation which others, such as Becketti, et. al. (1988), also did not include in their analysis. In order to get coefficients corrected for the selectivity bias of labor force participation one could estimate the model using a three-step approach as in Zabel (1994).

$$a^{*}_{it} = \alpha_o^{\prime} w_{it} + \mu_{it} \qquad \textbf{(10)}$$

respectively. The notation PV^{*}_{it} represents the latent variable on the tendency to be poor of individual i at time t. Assuming the error terms have a bivariate normal distribution, the coefficients of the poverty status and the attrition equations can be jointly estimated with the maximum likelihood method. A significant correlation coefficient between the error terms in the two equations provides a test for the selectivity bias from attrition.[15],[16]

As shown in Table 9, we found that there was a significant nonrandom effect of attrition on poverty status (the correlation term between the error terms (rho) was significant). The variables were significant and had the expected signs. The likelihood of being poor was not related to age, but females were more likely to be poor, while married persons and those with higher educational attainment were less likely to be poor. In addition, Blacks, Hispanics and the disabled are more likely to be poor than Whites, non-Hispanics, and non-disabled, respectively. All regions and metropolitan areas were less likely to be poor than the South and non-metropolitan areas.

To summarize, we used two methodologies to test for nonrandom effects of attrition on income and poverty estimates. First we examined the characteristics of stayers versus the original sample; second we tested for selectivity bias in structural equations during the life of the panel. We found some evidence of nonrandom attrition on income and poverty estimates. Specifically, for the initial status, stayers had some differences in terms of income determinants when compared to the original sample, but the differences were not consistent across income types. Stayers also differed in terms of initial poverty status. In addition, when we examined the effect of attrition over the life of the panel, we found some evidence of selectivity bias from nonrandom attrition on the determinants of income for persons below the poverty level and a significant effect on the determinants of poverty status.

Simulations of Poverty Status

In order to examine the magnitude of the effect that attrition may have on the overall poverty rate, we present some simulations of the poverty rate when information for persons missing one or more interviews was imputed. We used a longitudinal approach to impute missing data, specifically, reported information for the same person in another time period was used to impute information for the missing period. For example, if there was missing information in wave t, we used reported information for the same individual from previous (t-1) and later waves (t+1) to estimate the missing information. Using this information, we could obtain an estimate of the effect of attrition bias on poverty rates by calculating a poverty rate for the full sample (using imputed data for the missing periods) and

[15]The maximum likelihood estimation does not use the standard bivariate normal distribution because we do not observe the poverty status for cases that have attrited. See Van De Ven and Van Praag (1981) for the specific form of the distribution.

[16]We thank Jeffrey Zabel for providing the estimation programs for this technique.

comparing it to the poverty rate of respondents with complete information [Magnusson and Bergman (1990)].

For this paper, we used a straightforward approach to impute the missing information, specifically, family income and poverty threshold. If the person left the sample, we used the last reported information for the person for the missing months. If the missing interview was bounded by reported information, we carried forward the last reported information to a specific month (selected at random) and carried back later reported information for the remaining months with missing information. A similar approach was used by Lepkowski, Miller, and Luis (1993) to evaluate various methodologies to impute wave non-response in the SIPP. They compared the "carry-over" approach to a hot-deck approach and found that the carry-over approach tended to be more accurate in reporting monthly recipiency and amounts.[17] However, since the carryover approach did not change reported amounts, they found that for certain longitudinal measures, such as spells of recipiency, the carry-over imputations were biased while the hot-deck imputations were not.

Table 10 presents SIPP poverty estimates based on persons with fully reported data (SIPP), estimates based on all respondents including those with imputed information for persons with one or more missing interviews (SIPP-ATT)[18] , and poverty estimates from the CPS. We found that poverty estimates using the imputed data for cases with missing information were relatively close to the SIPP estimates using cases with complete reporting and were substantially lower than the CPS estimates. The SIPP-ATT estimate of the overall poverty rate was 11.0 percent, while the SIPP estimate was 10.5 percent and the CPS estimate was 13.5 percent. This result was consistent when we examined the estimates by characteristics such as age, race, sex, and education level. For example, the SIPP-ATT estimate for persons under 18 years of age was 18.1 percent compared to 17.1 and 20.6 percent for the SIPP and CPS estimates, respectively; the estimate for Whites was 8.3 percent compared to 7.7 and 10.7 percent; and the estimate for males was 9.3 percent compared to 8.5 and 11.7 percent. That is, we found that estimates with a correction for attrition explained a relatively small portion, about one-sixth, of the difference between the SIPP and the CPS estimates.

CONCLUSION

[17]The approach used by Lepkowski, Miller, and Luis (1993) was to blank out reported information for respondents with complete information in the 1987 Panel of SIPP and impute the “missing” information using alternative imputation strategies. They evaluated the imputation methods by comparing imputed and reported information as well as derived statistics.

[18]The universe for these estimates was persons with an interview in the first wave. Since the calendar year weights for SIPP include only persons with complete information during the year and excludes those with missing information, we weighted the SIPP-ATT estimates using the cross-sectional weight for January 1990.

In our examination of the differences in poverty rates between the SIPP and the CPS, we found large differences between the survey estimates, with SIPP poverty rates significantly lower than the CPS estimates. We found that the conceptual differences between the surveys represented a small part of the difference. Specifically, differences in the treatment of family composition accounted for one-sixth of the difference, and differences in the treatment of self-employment income had a very small effect on the estimates.

We concentrated our analysis on the effect of attrition on income and poverty estimates. We used two approaches to test for and estimate the effect of attrition. We compared the determinants of income and poverty at the start of the panel between respondents who remained in the panel and the original sample. We found some differences between stayers and the original sample in terms of total income and poverty status at the start of the panel. We also used a selectivity bias approach to examine the effect of attrition on income and poverty determinants during the life of the panel. We found some effect of attrition on income, but the effect was not consistent across income types, and we found a significant effect of attrition on determinants of poverty status.

To examine the potential magnitude of the effect of attrition on poverty estimates, we discussed some simulations of poverty status for the respondents missing one or more waves and compared them to poverty estimates for those with complete information. We found that attrition potentially explained a relatively small portion of the difference, about one-sixth of the difference between the SIPP and the CPS estimates. Separate calculations indicated that the effect of changes in family composition available for the SIPP but not the CPS are also responsible for approximately one-sixth of the difference. Assuming these two effects are independent, we conclude that the approximately one-third of the difference between the two estimates is attributable to compositional changes and attrition.

Based on the results described in this paper, we conclude that although attrition had an effect on income and poverty estimates in SIPP, the observed difference in the poverty estimates from SIPP and CPS do not appear to be the result of either attrition or the other methodological differences we examined. The differences in the estimates may be the result of better reporting of income at the lower end of the distribution, specifically, better reporting of recipiency of means-tested income and other short term spells of income recipiency. Further work in this area is warranted.

References

Becketti, S., W. Gould, L. Lillard, and F. Welch, 1988. "The Panel Study After Fourteen Years: An Evaluation." Journal of Labor Economics, 6, 472-492.

Bergman, L. and D. Magnusson, 1990. "General Issues About Data Quality in Longitudinal Research," In Data Quality in Longitudinal Research, eds. D. Magnusson and L. Bergman, Cambridge University Press.

Coder, J., D. Burkhead, A. Feldman-Harkins, and J. McNeil, 1987. "Preliminary Data from the SIPP 1983-84 Longitudinal Research File," Survey of Income and Program Participation Working Paper Series, No. 8702.

Folsom, R. E. and M. B. Witt, 1993. "Testing New Attrition Nonresponse Adjustment for SIPP," working paper, Research Triangle Institute.

Hausman J. A. and D. A. Wise, 1979. "Attrition Bias in Experimental and Panel Data: the Gary Income Maintenance Experiment," Econometrica, 47, 455-472.

Heckman, J., 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for such Models," Annals of Economic and Social Measurement, 5, 475-492.

----------------, 1979. "Sample Selection Bias as a Specification Error," Econometrica, 47, 153-161.

Hill, D., 1993. "An Investigation of 'Under' Estimates in the SIPP," working paper, The Survey Research Institute, The University of Toledo.

Jabine, T., 1990, Survey of Income and Program Participation Quality Profile, U.S. Government Printing Office, Washington, D.C.

Lepkowski, J., D. Miller, E. Luis, 1993. "Imputation for Wave Non-response in the SIPP," working paper, Institute for Social Research, University of Michigan.

Maddala, G.S., 1983. Limited-Dependent and Qualitative Variables in Econometrics, Econometric Society Monographs in Quantitative Economics, Cambridge University Press.

Olsen, R. J., 1982. "Distributional Tests for Selectivity Bias and a More Robust Likelihood Estimator," International Economic Review, Vol 23, No. 1, 223-240.

Ridder, G., 1990. "Attrition in Multi-Wave Panel Data." In Panel Data and Labor Market Studies, eds. J. Hartog, G. Ridder, and J. Theeuwes. Amsterdam: North-Holland.

Ruggles, P., 1990. Drawing the Line: Alternative Poverty Measures and Their Implications for Public Policy, Washington, D.C.: The Urban Institute Press.

Ruggles, P., and R. Williams, 1989. "Longitudinal Measures of Poverty." Review of Income and Wealth, 35, (September): 225-244.

Short, K., and E. McArthur, 1986. "Life Events and Sample Attrition in the Survey of Income and Program Participation," paper presented at the 1986 Annual Meetings of the American Statistical Association.

U.S. Bureau of the Census, 1991. "Survey of Income and Program Participation Users' Guide," Washington, D.C.

----------------, 1993. "Poverty in the United States: 1992," Current Population Reports, Series P60-185, U.S. Government Printing Office, Washington, D.C.

----------------, 1994. "Evaluating the Quality of Income Data Collected in the Annual Supplement to the March Current Population Survey and the Survey of Income and Program Participation," forthcoming, Washington, D.C.

Van De Ven, W.P.P.M. and B.M.S. Van Praag, 1981. "The Demand for Deductibles in Private Health Insurance," Journal of Econometrics, 17, 229-252.

Vaughan, D., 1989. "Reflections on the Income Estimates from the Initial Panel of the Survey of Income and Program Participation," in Individuals and Families in Transition: Understanding Change Through Longitudinal Data. Papers presented at the Social Science Research Council Conference in Annapolis, Maryland, March 16-18 1988, U.S. Bureau of the Census.

Weinberg, D., and E. Lamas, 1994. "Measurement of Need: The U.S. Poverty Line," 1993 Proceedings of the American Statistical Association, Social Statistics Section.

Zabel, J. E., 1994. "An Analysis of Attrition in the PSID and SIPP with an Application to Models of Labor Market Behavior," working paper, Tufts University.

Table 1. Percent of Persons in Poverty in 1990: SIPP and CPS

Characteristic	SIPP	SIPP-CS	CPS
All persons	10.5	11.1	13.5
AGE			
Under 18 years old	17.1	17.3	20.6
18 to 24 years old	11.3	13.2	15.9
25 to 34 years old	9.3	9.9	12.1
35 to 44 years old	6.2	6.8	8.5
45 to 54 years old	6.0	6.5	7.8
55 to 59 years old	7.4	7.0	9.0
60 to 64 years old	7.0	8.8	10.3
65 years old and over	8.3	9.2	12.2
RACE AND SPANISH ORIGIN			
White	7.7	8.7	10.7
Black	27.7	26.6	31.9
Other	15.6	13.7	-
Hispanic origin	22.4	21.8	28.1
Not Hispanic origin	9.4	10.1	-
SEX			
Male	8.5	9.2	11.7
Female	12.3	12.9	15.2

EDUCATION (25 years and older)

Less than high school	17.9	18.8	23.6
High school, no college	6.3	7.2	8.9
Some college, no degree	4.5	5.0	5.8
College degree	1.9	2.3	2.8

Table 2. Selected Characteristics of Attritors, Non-Attritors, and Individuals Who Left the Universe: 1990 Panel

Characteristic	Non-Attritors		Attritors		Left the Universe	
	Number	Percent	Number	Percent	Number	Percent
Total	41,939	100.0	14,350	100.0	1,860	100.0
AGE						
Less than 15 years	10,594	25.3	3,090	21.5	199	10.7
15 to 17 years	1,651	3.9	823	5.7	88	4.7
18 to 24 years	3,526	8.4	2,224	15.5	228	12.3
25 to 34 years	7,034	16.8	2,668	18.6	250	13.4
35 to 44 years	6,389	15.2	2,125	14.8	141	7.6
45 to 54 years	4,330	10.3	1,403	9.8	83	4.5
55 to 64 years	3,594	8.6	978	6.8	147	7.9
65 years and over	4,821	11.5	1,039	7.2	724	38.9
Chi-square statistic	917.46*				1768.84*	
SEX						
Male	19,556	46.6	7,225	50.3	1,012	54.4
Female	22,383	53.4	7,125	49.7	848	45.6
Chi-square statistic	59.28*				10.86*	
RACE						
White	35,423	84.5	10,772	75.1	1,517	81.6
Black	5,143	12.3	2,976	20.7	246	13.2
Other	1,373	3.3	602	4.2	97	5.2
Chi-square statistic	673.40*				59.89*	
HISPANIC ORIGIN						
Hispanic	4,048	9.7	2,148	15.0	403	21.7
Non-Hispanic	37,891	90.3	12,202	85.0	1,457	78.3
Chi-square statistic	308.50*				55.71*	

MARITAL STATUS						
Married	18,419	43.9	5,159	36.0	699	37.6
Widowed, divorced, or Separated	5,693	13.6	2,024	14.1	547	29.4
Never married	17,827	42.5	7,167	49.9	614	33.0
Chi-square statistic	297.41*				342.72*	
EDUCATION (PERSONS 18+)						
Total	29,694	100.0	10,437	100.0	1,573	100.0
Less than high school	6,399	21.5	2,497	23.9	657	41.8
High school	10,173	34.3	3,703	35.5	465	29.6
College	13,122	44.2	4,237	40.6	451	28.7
Chi-square statistic	46.04*				229.31*	
FAMILY INCOME						
Zero or negative income	433	1.0	375	2.6	39	2.1
$1 to $999	6,173	14.7	2,622	18.3	514	27.6
$1,000 to $1,999	8,772	20.9	3,293	22.9	561	30.2
$2,000 to $2,999	8,094	19.3	2,715	18.9	305	16.4
$3,000 to $3,999	6,508	15.5	1,834	12.8	200	10.8
$4,000 to $4,999	4,522	10.8	1,234	8.6	78	4.2
$5,000 to $5,999	2,686	6.4	793	5.5	52	2.8
$6,000 to $6,999	1,608	3.8	495	3.4	33	1.8
$7,000 to $7,999	904	2.2	271	1.9	19	1.0
$8,000 to $8,999	638	1.5	194	1.4	14	0.8
$9,000 to $9,999	371	0.9	135	0.9	8	0.4
$10,000 and over	1,232	2.9	389	2.7	37	2.0
Chi-square statistic	423.62*				220.26*	

Table 2. Selected Characteristics of Attritors, Non-Attritors, and Individuals who left the Universe: 1990 Panel

Characteristic	Non-Attritors		Attritors		Left the Universe	
	Number	Percent	Number	Percent	Number	Percent
POVERTY STATUS						
Poor	5,634	13.4	2,461	17.1	342	18.4
Non-poor	36,305	86.6	11,709	81.6	1,518	81.6
Chi-square statistic	210.64*				0.00	
POVERTY THRESHOLD						
$1 to $999	22,873	54.5	7,945	55.4	1,323	71.1
$1,000 to $1,999	18,684	44.6	6,221	43.4	516	27.7
$2,000 to $2,999	382	0.9	184	1.3	21	1.1
Chi-square statistic	19.47*				168.40*	
EMPLOYMENT STATUS						
Total	31,529	100.0	11,339	100.0	1,667	100.0
Employed all month	18,982	60.2	6,744	59.5	535	32.1
With job part of month	740	2.3	349	3.1	37	2.2
Unemployed	868	2.8	603	5.3	58	3.5
Not in labor force	10,939	34.7	3,643	32.1	1,037	62.2
Chi-square statistic	194.32*				573.42*	

* Significant at the .05 level.

Table 3. Probit Regressions for Attrition: All Persons 15 and over and Persons Below Poverty Level

Independent variables	All persons	Below poverty level
Constant	-0.447* (0.223)	5.691* (2.172)
Age	0.008 (0.005)	-0.062 (0.048)
Age squared	-0.00003 (0.00005)	0.0005 (0.0004)
Education	0.015* (0.002)	-0.017* (0.026)
Change in education	-0.404* (0.04)	-0.013 (0.650)
Income	0.00002* (8.1E-6)	-
Income squared	-8.6E-11 (6.5E-11)	-
Female	-0.185* (0.026)	0.213 (0.349)
Married	-1.101* (0.069)	-0.603 (1.057)
Change in marriage	-0.744* (0.069)	0.373 (1.130)
Widowed, divorced, or separated	-0.829* (0.079)	-0.886 (0.760)

Change in widowed, divorced, or separated	-0.532* (0.079)	-0.524 (0.759)
Employed full time	0.276* (0.048)	0.239 (0.468)
Change in employed full time	0.319* (0.042)	1.636* (0.446)
Employed part time	0.163* (0.074)	-
Change in employed part time	-0.612* (0.027)	-
Unemployed	0.046 (0.082)	-0.466 (0.548)
Change in unemployed	-0.502* (0.186)	-
Black	0.636* (0.035)	1.265* (0.273)
Hispanic	0.339* (0.041)	0.688* (0.312)
Disabled	0.351* (0.036)	0.163 (0.310)
Mover	1.916* (0.027)	0.161 (0.294)
Nonrelatives	0.486* (0.065)	0.061 (6.110)
Observations	32,478	2,749

Standard errors shown in parentheses.
* Significant at .05 level.

Table 4. Initial Income Regressions for Original and Remaining Sample: All Persons

Independent variables	All income: Original sample (month 1)	All income: Remaining sample (month 32)	All income: Difference	Labor income: Original sample (month 1)	Labor income: Remaining sample (month 32)	Labor income: Difference	Nonlabor income, Total: Original sample (month 1)	Nonlabor income, Total: Remaining sample (month 32)	Nonlabor income, Total: Difference	Nonlabor income, Means-Tested: Original sample (month 1)	Nonlabor income, Means-Tested: Remaining sample (month 32)	Nonlabor income, Means-Tested: Difference
Constant	3.525* (0.053)	3.358* (0.059)	0.167* (0.079)	4.403* (0.051)	4.391* (0.056)	0.012 (0.076)	0.745* (0.080)	0.780* (0.089)	-0.035 (0.120)	5.173* (0.101)	5.165* (0.115)	0.008 (0.153)
Age	0.125* (0.002)	0.134* (0.003)	-0.009* (0.004)	0.116* (0.003)	0.118* (0.003)	-0.002 (0.004)	0.059* (0.003)	0.052* (0.004)	0.007 (0.005)	0.022* (0.004)	0.022* (0.005)	0.000 (0.006)
Age square	-0.001* (0.00002)	-0.001* (0.00002)	0.000 (0.000)	-0.001* (0.00003)	-0.001* (0.00004)	0.000 (0.000)	0.0002* (0.00003)	0.0003* (0.00004)	-0.000 (0.000)	-0.0001* (0.00004)	-0.0001* (0.00005)	0.000 (0.000)
Education	0.045* (0.001)	0.046* (0.001)	-0.001 (0.001)	0.030* (0.001)	0.029* (0.001)	0.001 (0.001)	0.024* (0.002)	0.027* (0.002)	-0.003 (0.003)	0.010* (0.002)	0.010* (0.003)	0.000 (0.003)
Female	-0.836* (0.015)	-0.854* (0.016)	0.018 (0.022)	-0.491* (0.009)	-0.515* (0.010)	0.024 (0.013)	-0.007 (0.022)	0.002 (0.024)	-0.009 (0.033)	-0.102* (0.030)	-0.126* (0.035)	0.024 (0.046)
Married	-0.230* (0.016)	-0.237* (0.018)	0.007 (0.024)	0.080* (0.010)	0.072* (0.011)	0.008 (0.015)	-0.591* (0.024)	-0.569* (0.026)	-0.022 (0.035)	-0.116* (0.031)	-0.146* (0.035)	0.030 (0.047)
Black	-0.003 (0.023)	-0.015 (0.026)	0.012 (0.035)	-0.115* (0.014)	-0.124* (0.016)	0.009 (0.021)	0.013 (0.036)	-0.001 (0.041)	0.014 (0.055)	0.051 (0.029)	0.051 (0.034)	0.000 (0.045)
Hispanic	-0.033 (0.026)	-0.018 (0.029)	-0.015 (0.039)	-0.150* (0.016)	-0.140* (0.018)	-0.010 (0.024)	-0.188* (0.042)	-0.142* (0.047)	-0.046 (0.063)	-0.011 (0.036)	-0.010 (0.041)	-0.001 (0.055)
Disabled	-0.438* (0.021)	-0.439* (0.023)	0.000 (0.027)	-0.280* (0.015)	-0.286* (0.016)	0.000 (0.027)	0.754* (0.030)	0.722* (0.033)	0.000 (0.027)	0.084* (0.028)	0.104* (0.032)	0.000 (0.027)
Northeast	0.015 (0.020)	0.003 (0.022)	0.012 (0.030)	0.134* (0.013)	0.123* (0.014)	0.011 (0.019)	0.028 (0.030)	0.023 (0.033)	0.005 (0.045)	0.147* (0.036)	0.135* (0.041)	0.012 (0.055)
Midwest	0.011 (0.019)	0.004 (0.021)	0.000 (0.027)	0.021 (0.012)	0.018 (0.013)	0.000 (0.027)	-0.058* (0.029)	-0.076* (0.031)	0.000 (0.027)	0.087* (0.035)	0.096* (0.039)	0.000 (0.027)
West	0.060* (0.021)	0.054* (0.023)	0.006 (0.031)	0.099* (0.013)	0.103* (0.014)	-0.004 (0.019)	0.110* (0.031)	0.069* (0.034)	0.041 (0.046)	0.323* (0.037)	0.329* (0.042)	-0.006 (0.056)
Metropolitan Residence	0.168* (0.019)	0.168* (0.021)	0.000 (0.027)	0.201* (0.012)	0.209* (0.013)	0.000 (0.027)	0.051 (0.028)	0.064* (0.03)	0.000 (0.027)	0.054 (0.031)	0.057 (0.035)	0.000 (0.027)

R-squared	0.190	0.199		---	0.288	0.292		---	0.381	0.381		---	0.098	0.107		---
Observations	39,996	33,025		6,971	25,055	20,928		4,127	32,136	27,006		5,130	3,396	2,674		722
F-tests:																
All coefficients	0.99				0.99				1.07*				0.94			
All but constant	0..98				0.96				1.03*				0.93			

Standard errors shown in parentheses.
* Significant at .05 level.

Table 5. Initial Income Regressions for Original and Remaining Sample: Persons Below Poverty

Independent variables	All income			Labor income			Nonlabor income: Total			Nonlabor income: Means-Tested		
	Original sample (month 1)	Remaining sample (month 32)	Difference	Original sample (month 1)	Remaining sample (month 32)	Difference	Original sample (month 1)	Remaining sample (month 32)	Difference	Original sample (month 1)	Remaining sample (month 32)	Difference
Constant	5.427* (0.174)	5.353* (0.202)	0.074 (0.267)	5.049* (0.206)	5.134* (0.243)	-0.085 (0.319)	3.505* (0.241)	3.604* (0.273)	-0.099 (0.364)	5.582* (0.115)	5.604* (0.126)	-0.022 (0.171)
Age	0.020* (0.007)	0.023* (0.009)	-0.003 (0.011)	0.062* (0.011)	0.059* (0.013)	0.003 (0.017)	0.041* (0.010)	0.035* (0.011)	0.006 (0.015)	0.015* (0.005)	0.015* (0.005)	0.000 (0.007)
Age square	0.00001 (0.0001)	-0.0002* (0.0001)	0.0002* (0.0001)	-0.001* (0.0001)	-0.001* (0.0002)	-0.0001 (0.0002)	-0.0001 (0.0001)	-0.0001 (0.0001)	0.0000 (0.0001)	-0.0002* (0.0001)	-0.0002* (0.0001)	0.0000 (0.0001)
Education	-0.041* (0.004)	-0.040* (0.005)	-0.001 (0.006)	-0.008* (0.004)	-0.009* (0.004)	0.001 (0.005)	-0.054* (0.006)	-0.052* (0.007)	-0.002 (0.009)	-0.002 (0.003)	-0.003 (0.003)	0.001 (0.004)
Female	-0.039 (0.054)	-0.019 (0.062)	-0.020 (0.082)	-0.206* (0.043)	-0.246* (0.049)	0.040 (0.065)	0.394* (0.073)	0.386* (0.084)	0.008 (0.111)	0.089* (0.038)	0.072 (0.043)	0.017 (0.057)
Married	-0.365* (0.058)	-0.377* (0.065)	0.012 (0.087)	0.351* (0.046)	0.351* (0.054)	0.000 (0.071)	-0.993* (0.078)	-0.955* (0.087)	-0.038 (0.117)	-0.099* (0.040)	-0.071 (0.044)	-0.028 (0.059)
Black	0.298* (0.058)	0.281* (0.066)	0.0170 (0.0879)	0.276* (0.051)	0.277* (0.059)	-0.001 (0.078)	0.433* (0.076)	0.360* (0.086)	0.073 (0.115)	0.035 (0.030)	0.040 (0.033)	-0.005 (0.045)
Hispanic	0.193* (0.068)	0.195* (0.077)	-0.002 (0.103)	0.106 (0.055)	0.059 (0.065)	0.047 (0.085)	0.019 (0.093)	0.106 (0.105)	-0.087 (0.140)	0.072 (0.038)	0.031 (0.042)	0.041 (0.057)
Disabled	-0.029 (0.058)	-0.029 (0.064)	0.000 (0.086)	-0.184* (0.055)	-0.165* (0.063)	-0.019 (0.084)	0.433* (0.076)	0.457* (0.084)	-0.024 (0.113)	-0.043 (0.032)	-0.032 (0.035)	-0.011 (0.047)
Northeast	0.073 (0.070)	0.055 (0.078)	0.018 (0.105)	0.100 (0.066)	0.122* (0.076)	-0.022 (0.101)	0.388* (0.091)	0.373* (0.101)	0.015 (0.136)	0.199* (0.039)	0.218* (0.043)	-0.019 (0.058)
Midwest	0.111 (0.064)	0.105 (0.072)	0.006 (0.096)	0.004 (0.053)	-0.017 (0.060)	0.021 (0.080)	0.146 (0.084)	0.139 (0.094)	0.007 (0.126)	0.130* (0.037)	0.139* (0.041)	-0.009 (0.055)
West	0.078 (0.072)	0.119 (0.083)	-0.041 (0.110)	0.173* (0.056)	0.184* (0.067)	-0.011 (0.087)	0.249* (0.097)	0.303* (0.112)	-0.054 (0.148)	0.308* (0.047)	0.335* (0.052)	-0.027 (0.070)
Metropolitan Residence	-0.013 (0.058)	-0.029 (0.064)	0.016 (0.086)	-0.049 (0.046)	-0.045 (0.052)	-0.004 (0.069)	0.085 (0.077)	0.088 (0.085	-0.003 (0.115)	0.028 (0.034)	0.016 (0.037)	0.012 (0.050)

R squared	0.068	0.069	---	0.191	0.200	---	0.247	0.244	---	0.067	0.065	---
Observations	3,777	2,911	866	1,144	863	281	3,041	2,384	657	1,559	1,250	309
F-tests:												
All coefficients	1.09*			1.03			1.04*			1.18*		
All but constant	1.08*			1.03			1.07*			1.13		

Standard errors shown in parentheses.
* Significant at .05 level.

Table 6. Initial Poverty Status Regressions for Attritors and Nonattritors

Poverty status

Independent variables	All persons	Attritors	Non-Attritors	Differences
Constant	0.034 (0.084)	0.145 (0.149)	-0.018 (0.102)	0.163 (0.181)
Age	0.002 (0.002)	-0.001 (0.005)	0.003 (0.003)	-0.004 (0.006)
Age square	0.0004* (0.00003)	0.0001 (0.0001)	0.00003 (0.00004)	0.0001 (0.0001)
Education	-0.312* (0.022)	-0.331* (0.038)	-0.319* (0.025)	-0.012 (0.045)
Female	0.217* (0.017)	0.186* (0.031)	0.232* (0.021)	-0.046 (0.037)
Married	-0.504* (0.019)	-0.440* (0.036)	-0.519* (0.023)	0.079 (0.043)
Black	0.443* (0.022)	0.336* (0.037)	0.454* (0.028)	-0.118* (0.046)
Hispanic	0.411* (0.025)	0.237* (0.043)	0.489* (0.031)	-0.252* (0.053)
Disabled	0.323* (0.022)	0.233* (0.039)	0.358* (0.027)	-0.125* (0.047)
Northeast	-0.166* (0.024)	-0.154* (0.043)	-0.163* (0.029)	0.009 (0.052)
Midwest	-0.123* (-0.023)	-0.050 (0.044)	-0.138* (0.027)	0.088 (0.052)
West	-0.130* (0.024)	-0.035 (0.042)	-0.172* (0.030)	0.137* (0.052)
Metropolitan Residence	-0.233* (0.021)	-0.146* (0.043)	-0.264* (0.024)	0.118* (0.049)
Loglikelihood value	-14,093.89	-4,220.73	-9,838.13	-
Observations	44,266	11,260	33,006	-
Likelihood-ratio test		70.06*		

Standard errors shown in parentheses.

* Significant ar .05 level.

Table 7. The Effect of Attrition on the Coefficient Estimates of Income Determinants for Individuals 15 Years of Age or Older

All persons	With labor income	With non-labor income

		Income equation			Income equation			Total			Means-Tested	
Independent variables	Attrition equation	With lambda	No lambada	Attrition equation	With lambada	No lambada	Attrition equation	With lambada	No lambada	Attrition equation	With lambada	No lambada
Constant	-0.9852* (0.187)	0.3774* (0.057)	0.3597* (0.055)	-2.0753* (0.295)	-0.1807* (0.056)	-0.1982* (0.054)	-0.3247 (0.273)	-3.3681* (0.113)	-3.742* (0.111)	-0.9303 (0.693)	1.1751* (0.123)	1.1659* (0.121)
Age	-0.0134* (0.005)	0.0681* (0.002)	0.0683* (0.002)	0.0317* (0.011)	0.1007* (0.003)	0.1011* (0.003)	-0.0257* (0.007)	0.0691* (0.004)	0.0691* (0.004)	-0.0096 (0.019)	-0.005 (0.005)	-0.005 (0.005)
Age square	0.0001 (0.0001)	-0.0006* (0.0000)	-0.0006* (0.0000)	-0.0004* (0.0001)	-0.0011* (0.0000)	-0.0011* (0.0000)	0.0002* (0.0001)	-0.0000 (0.0000)	-0.0000 (0.0000)	0.0000 (0.0002)	-0.0001* (0.0000)	-0.0001* (0.0000)
Education	0.0191* (0.002)	0.0354* (0.001)	0.0355* (0.001)	0.0285* (0.003)	0.0355* (0.001)	0.0356* (0.001)	0.0154* (0.003)	0.0095* (0.002)	0.0095* (0.002)	-0.0176 (0.009)	0.0108* (0.002)	0.0107* (0.002)
Female	-0.1926* (0.030)	-0.7556* (0.013)	-0.7383* (0.013)	-0.1890* (0.043)	-0.4028* (0.010)	-0.4014* (0.010)	-0.1707* (0.042)	-0.1338* (0.024)	-0.1333* (0.024)	-0.4037* (0.133)	-0.0431 (0.032)	-0.0403 (0.032)
Married	-0.1573* (0.032)	-0.1782* (0.014)	-0.1756* (0.014)	-0.2241* (0.046)	0.0731* (0.011)	0.0754* (0.011)	-0.0525 (0.045)	-0.4941* (0.026)	-0.4934* (0.026)	0.0949 (0.139)	0.0873* (0.033)	0.0877* (0.033)
Black	0.5528* (0.044)	-0.092* (0.021)	-0.0975* (0.021)	0.7709* (0.061)	-0.1463* (0.016)	-0.1522* (0.016)	0.4321* (0.065)	0.1124* (0.041)	0.1107* (0.041)	0.3728* (0.124)	0.0010 (0.029)	0.0012 (0.029)
Hispanic	0.1887* (0.051)	-0.0988* (0.024)	-0.1015* (0.024)	0.2025* (0.070)	-0.1441* (0.018)	-0.1462* (0.018)	0.1823* (0.078)	-0.0722 (0.051)	-0.0731 (0.049)	0.1115 (0.155)	-0.0422 (0.038)	-0.0422 (0.038)
Disabled	0.2799* (0.042)	-0.3880* (0.019)	-0.3902* (0.019)	0.2504* (0.081)	-0.2220* (0.019)	-0.2232* (0.019)	0.1267* (0.055)	0.7664* (0.032)	0.7659* (0.032)	-0.1181 (0.122)	0.0094 (0.029)	0.0101 (0.029)
Northeast	0.0385 (0.040)	0.0114 (0.018)	0.0119 (0.019)	0.0352 (0.057)	0.1187* (0.014)	0.1197* (0.014)	-0.0177 (0.055)	-0.0293 (0.033)	-0.0288 (0.033)	0.3131* (0.151)	0.0999* (0.036)	0.0983* (0.036)
Midwest	-0.4616* (0.041)	0.0108 (0.017)	0.0138 (0.017)	-0.5286* (0.058)	0.0365* (0.013)	0.0386* (0.013)	-0.4866* (0.057)	-0.0349 (0.032)	-0.0335 (0.032)	0.1004 (0.151)	0.0945* (0.036)	0.0939* (0.036)
West	-0.0563 (0.041)	0.0436* (0.018)	0.0437* (0.018)	-0.1038 (0.058)	0.0875* (0.014)	0.0878* (0.014)	-0.0432 (0.056)	0.0838* (0.034)	0.0839* (0.039)	0.3188* (0.156)	0.2524* (0.038)	0.2498* (0.038)
Metropolitan Residence	0.2506* (0.041)	0.1439* (0.017)	0.1417* (0.017)	0.2393* (0.059)	0.1914* (0.013)	0.1899* (0.013)	0.2074* (0.057)	0.0256 (0.031)	0.0246 (0.031)	0.4759* (0.150)	0.0544 (0.033)	0.0518 (0.032)
Lambda		-0.0186 (0.016)			-0.0105 (0.009)			-0.0084 (0.030)			-0.0135 (0.027)	
Mover	2.1529* (0.032)			2.2774* (0.046)			2.5735* (0.044)			2.3864* (0.118)		

Nonrelatives	0.602*			0.6241*			0.550*			0.5322		
	(0.081)			(0.103)			(0.133)			(0.299)		
Observations	32,551			15,949			20,117			2,436		
R squared		0.169	0.170		0.309	0.309		0.380	0.380		0.055	0.055

Standard errors shown in parentheses.
* Significant at .05 level.

Table 8. The Effect of Attrition on the Coefficient Estimates of Income Determinants for Poor Individuals 15 Years of Age or Older

	All persons		With labor income			With non-labor income
		Income equation		Income equation	Total	Means-Tested

Independent variables	Attrition equation	With lambda	No lambada	Attrition equation	With lambada	No lambada	Attrition equation	With lambada	No lambada	Attrition equation	With lambada	No lambada
Constant	-0.441	1.168*	1.043*	-8.714	0.102	-0.141	0.597	0.436	0.373	0.485	1.295*	1.289*
	(0.995)	(0.210)	(0.210)	(4.615)	(0.811)	(0.706)	(1.083)	(0.259)	(0.259)	(1.338)	(0.146)	(0.146)
Age	-0.037	-0.001	-0.002	0.129	0.068	0.079	-0.059	0.011	0.104	-0.066	-0.001	-0.001
	(0.029)	(0.008)	(0.008)	(0.206)	(0.045)	(0.041)	(0.032)	(0.009)	(0.009)	(0.041)	(0.006)	(0.006)
Age square	0.0004	-0.00000	-0.00000	-0.0003	-0.001	-0.001	0.001	0.0000	0.0000	0.0007	-0.00001	-0.00001
	(0.0003)	(0.0001)	(0.0001)	(0.003)	(0.001)	(0.001)	(0.0003)	(0.0001)	(0.0001)	(0.0004)	(0.0001)	(0.0001)
Education	-0.044*	-0.018*	-0.024*	0.026	-0.021	-0.021	-0.051*	-0.017*	-0.025*	-0.029	0.001	0.001
	(0.018)	(0.005)	(0.005)	(0.067)	(0.016)	(0.016)	(0.019)	(0.006)	(0.006)	(0.022)	(0.004)	(0.004)
Female	-0.538*	-0.123	-0.039	-0.144	-0.226	-0.202	-0.639*	0.008	0.111	-0.378	0.159*	0.168*
	(0.231)	(0.069)	(0.066)	(0.769)	(0.175)	(0.171)	(0.254)	(0.085)	(0.081)	(0.327)	(0.048)	(0.048)
Married	0.129	-0.138	-0.136	-1.164	-0.028	0.044	0.288	-0.421*	-0.446*	0.042	0.214*	0.213*
	(0.261)	(0.075)	(0.076)	(0.706)	(0.174)	(0.171)	(0.289)	(0.095)	(0.095)	(0.379)	(0.057)	(0.057)
Black	0.618*	0.092	0.052	-0.658	0.267	0.271	0.684*	0.185*	0.134*	0.892*	0.006	0.007
	(0.189)	(0.054)	(0.053)	(0.717)	(0.165)	(0.165)	(0.201)	(0.064)	(0.064)	(0.231)	(0.036)	(0.034)
Hispanic	0.018	0.079	0.093	-0.781	0.267	0.312	0.024	0.119	0.139	0.058	-0.002	-0.0001
	(0.232)	(0.068)	(0.068)	(0.759)	(0.165)	(0.174)	(0.254)	(0.083)	(0.084)	(0.284)	(0.043)	(0.043)
Disabled	-0.325	0.061	0.082	0.357	-0.079	-0.059	-0.347	0.208*	0.237*	-0.122	-0.095*	-0.093*
	(0.205)	(0.059)	(0.059)	(1.049)	(0.258)	(0.254)	(0.218)	(0.069)	(0.069)	(0.246)	(0.038)	(0.037)
Northeast	0.101	0.105	0.101	2.305	0.152	0.115	-0.081	0.252*	0.253*	0.119	0.062	0.059
	(0.240)	(0.068)	(0.068)	(1.265)	(0.247)	(0.239)	(0.259)	(0.082)	(0.082)	(0.295)	(0.045)	(0.045)
Midwest	-0.173	0.089	0.109	1.715	0.119	0.098	-0.378	0.184*	0.219*	-0.069	0.053	0.053
	(0.227)	(0.065)	(0.065)	(1.005)	(0.216)	(0.212)	(0.248)	(0.079)	(0.079)	(0.281)	(0.043)	(0.043)
West	0.342	0.255*	0.218*	1.511	0.237	0.201	0.137	0.343*	0.317*	-0.263	0.241*	0.236*
	(0.268)	(0.083)	(0.083)	(1.016)	(0.219)	(0.211)	(0.291)	(0.101)	(0.101)	(0.347)	(0.056)	(0.056)
Metropolitan Residence	0.491*	0.091	0.054	1.198	0.041	0.032	0.655*	0.064	0.017	0.836*	0.018	0.009
	(0.229)	(0.062)	(0.062)	(0.741)	(0.171)	(0.169)	(0.251)	(0.075)	(0.075)	(0.312)	(0.042)	(0.042)
Lambda		-0.089*			-0.001			-0.097*			-0.024	
		(0.021)			(0.001)			(0.024)			(0.024)	
Mover	2..979*			-				3.047*		2.779*		
	(0.203)							(0.217)		(0.249)		
Nonrelatives	0.757*			-			0.601			0.201		
	(0.351)						(0.377)			(0.499)		

Observations	1,069			80				970			724		
R squared		0.057	0.042			0.243	0.239		0.118	0.103		0.147	0.145

Standard errors shown in parentheses.
* Significant at .05 level.

Table 9. The Effect of Attrition on Determinants of Poverty Status

Independent variables	Attrition equation	Poverty equation

Constant	-0.745*	0.998*
	(0.087)	(0.111)
Age	0.003	0.0001
	(0.002)	(0.003)
Age square	-0.00001	0.00002
	(0.00002)	(0.00003)
Education	-0.112*	-0.384*
	(0.015)	(0.023)
Female	-0.125*	0.243*
	(0.014)	(0.020)
Married	-0.146*	-0.464*
	(0.016)	(0.022)
Black	0.343*	0.415*
	(0.021)	(0.028)
Hispanic	0.154*	0.494*
	(0.023)	(0.029)
Disabled	0.219*	0.326*
	(0.019)	(0.026)
Northeast	0.037*	-0.146*
	(0.020)	(0.028)
Midwest	-0.239*	-0.035*
	(0.016)	(0.026)

West	-0.019 (0.020)	-0.140* (0.029)
Metropolitan Residence	0.143* (0.189)	-0.258* (0.024)
Mover	1.129* (0.147)	
Nonrelatives	0.362* (0.037)	
Observations	42,698	
Rho		-0.5938* (0.046)

Standard errors shown in parentheses.
* Significant at .05 level.

Table 10. Percent of Persons in Poverty in 1990: SIPP and CPS

	SIPP	SIPP-ATT	CPS

Characteristic				
All Persons	10.5	11.0		13.5
AGE				
Under 18 years old	17.1	18.1		20.6
18 to 24 years old	11.3	12.6	15.9	
25 to 34 years old	9.3	9.7	12.1	
35 to 44 years old	6.2	7.1	8.5	
45 to 54 years old	6.0	6.1	7.8	
55 to 59 years old	7.4	7.3	9.0	
60 to 64 years old	7.0	7.0	10.3	
65 years old and over	8.3	8.5	12.2	
RACE AND SPANISH ORIGIN				
White	7.7	8.3	10.7	
Black	27.7	28.4	31.9	
Other	15.6	16.9	-	
Hispanic Origin	22.4	23.7	28.1	
Not Hispanic Origin	9.4	9.9	-	
SEX				
Male	8.5	9.3	11.7	
Female	12.3	12.7	15.2	
EDUCATION (25 years and older)				
Less than high school	17.9	17.8	23.6	
High school, no college	6.3	6.8	8.9	
Some college, no degree	4.5	4.9	5.8	

College degree	1.9	2.2	2.8