# Measurement Error and Poverty Rates of Widows

**Kathleen McGarry**

ABSTRACT

*Estimates of the poverty rate and of the probability of entering or exiting poverty are biased when income is observed with error. I estimate a variance components model of income which contains a white noise error term and then treat this component as an approximation of the error in observed income. By comparing poverty rates calculated with and without this estimated measurement error, I conclude that observation error causes the poverty rate to be overestimated around two percentage points on average. However, eliminating observation error substantially reduces the probability of transiting either into or out of poverty. These reductions imply that the amount of permanent poverty is underestimated when measurement error is ignored.*

## I. Introduction

In examining dynamic aspects of poverty, economists have focused their attention on panel data. By observing an individual over time, panel surveys allow one to construct a sequence of observations on income. These values can be used to derive an individual's poverty status in each period. Thus, it is possible to obtain estimates of the average length of a spell of poverty and of the extent to which poverty is permanent. The accuracy of these calculated measures of permanence, however, depends on the quality of the underlying measures of income. If an individual's income is observed with error, it is possible that her resulting poverty status will also be incorrect. This paper tests the importance of measurement error in dynamic studies of poverty by approximating it

*The author is an assistant professor of economics at the University of California, Los Angeles.* *She would like to thank James N. Brown, Michael D. Hurd and Audrey Light for their helpful comments. The data used in this article can be obtained beginning July 1995 through July 1998 from the author at the Department of Economics, University of California, Los Angeles, Los Angeles, CA 90024-1477.*
[Submitted June 1991; accepted October 1993]

with a white noise process. Data from the Retirement History Survey (RHS) are used to estimate a variance components model of income. The detailed error structure allows shocks to income from different sources to be examined separately. Estimates of the parameters of the model are used to derive predicted values of income in any given period. I compare measures of income calculated both with and without this residual error component. To the extent that observation error follows a white noise process, its influence on income, and thus spurious transitions into and out of poverty, are eliminated by this procedure. Since legitimate changes in income which last only one period cannot be distinguished from true measurement error, the impact of the latter is overstated. Such a procedure does, however, provide an indication of the importance of transitory shocks with respect to the calculation of various poverty rates.

## II. Previous Studies

The majority of the work examining income mobility and the dynamics of poverty focuses on the nonelderly population. Most of these studies (for example, Coe 1978, Duncan 1984, and Bane and Ellwood 1986) find a high probability of exiting from poverty. Because the focus of these studies is the population under age 65, the high exit rates from poverty are reasonable. For example, one could think of an unemployed person returning to work, or a single mother obtaining subsidized child care and entering the labor force. In fact, high exit rates would be expected if the many government programs aimed at alleviating poverty were successful. The situation for the elderly is expected to be different. Their incomes, and the incomes of widows in particular, are typically thought to be comprised of comparatively stable components. In 1986, for example, Social Security and pensions accounted for 54 percent of the income of the elderly (Social Security Administration 1988). If a sizable portion of income is from sources which adjust only for inflation, real income should be relatively constant. Surprisingly, despite the intuition, many studies have found a degree of income mobility among the elderly similar to that of the general population. Coe (1978) separately examines those over age 65. He finds 13.6 percent of the elderly population in poverty in a given year, but only 2.1 percent poor in all nine years of the study. Holden, Burkhauser, and Meyers (1986) study transitions into and out of poverty after retirement and widowhood. They find well over 50 percent of the widows exiting poverty after only a single year in that state, and even higher exit rates of close to 70 percent for couples. Many of these short, temporary spells for widows commence in the period immediately following their husbands' deaths. Although a period of financial adjustment after the death of a spouse is likely, improper measurement of income due to the survey procedures themselves causes much of this observed poverty (Burkhauser, Holden, and Myers 1986).[1]

1. In most surveys, including the RHS, respondents are asked to report only their own income for the preceding year. A recently widowed woman whose husband was alive for much of the previous year may have little of her own income to report. If this were the case, the recorded amount of income would not yield a true indication of the widow's financial status.

Hurd and Wise (1989) use the RHS to study the impact of widowhood on levels of wealth and income. They find a substantial amount of persistence of poverty; 82 percent of the widows who are poor in 1979, based on their total wealth, were also poor in 1969. However, when defining poverty status by income rather than wealth (as is done in calculating official poverty statistics), this figure is only 60 percent. Despite the interest in the amount of movement about the poverty line, little attention has been given to measuring this statistic accurately. While errors in observed income will approximately offset each other in aggregate estimates of the proportion poor, estimates of the amount of movement about the poverty line will likely be significantly biased by the existence of measurement error; more movement will be observed than actually occurs. Two papers in particular attempt to deal with the influence of observation error on calculations of the persistence of poverty. As a correction for measurement error, Bound et al. (1991) in their study of widows, ignore any transition into or out of poverty if the subsequent length of time spent in that state is limited to a single year. They require two periods of either poverty or nonpoverty before a transition to that state is considered valid. Bane and Ellwood (1986) use a similar procedure in their study of the nonelderly. They eliminate one period spells if the associated change in income is less than one-half of the ratio of income to the appropriate poverty line. These methods of correcting for measurement error while simple to implement are also likely to exclude a number of true transitions into and out of poverty. Even if there were no measurement error in reported income, one would expect to observe some individuals experiencing brief spells of poverty. The model presented here allows for the existence of one period spells even after correcting for measurement error. Both the raw data and the predictions generated by the model show that a significant fraction of the population is expected to spend a single period either in or out of poverty.

## III. The Data

The Retirement History Survey (RHS) interviewed respondents every two years from 1969 to 1979. The initial sample consists of men and unmarried women who were born between 1905 and 1911. In 1969 all primary respondents were between the ages of 58 and 63. Like Bound et al. I will concentrate on widows. Only those women who were widowed at the start of the survey and who were interviewed in each of the five subsequent periods are selected into the sample. I also require that the respondents remain widowed throughout the study. Thus, those women who drop out of the survey or remarry are excluded.[2] The resulting sample consists of 934 widows. By requiring that the woman be widowed at the start of the survey, the problem of obtaining an accurate measure of income in the first year of widowhood is avoided.[3] However, by focusing solely on these widows, I limit the sample to women who were widowed fairly early in

2. The incidence of remarriage is quite low among widows. Only about 4 percent of the women remarry over the sample period.

3. See Burkhauser, Holden and Myers (1986) for a discussion of the problem.

their lives. This group of younger widows may not be representative of women who become widowed at older ages. The composition of income changes substantially over the panel. At the start of the panel, many of these widows were employed. In 1969, 649 women reported nonzero earnings and, for the sample as a whole, labor income comprised 56 percent of total income, on average. By the end of the survey, only 20 percent reported positive earnings and, on average, earnings accounted for 8 percent of income. As the sample aged, an increasing number of the respondents became eligible for and chose to receive Social Security benefits.[4] In 1979, Social Security was the largest component of income. It accounted for 63 percent of income, on average—an increase of 44 percentage points from 1969. The establishment at the federal level of Supplemental Security Income (SSI) in 1974 further augmented the incomes of the poorest widows. Over time, the retirement of those with labor earnings and the receipt of Social Security benefits and SSI by those with little other income acted jointly to reduce the variance of the distribution. During the period covered by the RHS, the sample variance was reduced by approximately 35 percent, and the mean income of the sample declined almost monotonically from $7,289 in 1969 to $5,424 in 1979 (both figures in 1979 dollars). One might expect movements into poverty to be associated with a cessation of earnings. Similarly, the initial receipt of transfers from government programs would be expected to result in an increase in income, and perhaps a corresponding exit from poverty. While these assumptions are undoubtedly valid in many cases, a large number of transitions are unexplained by these factors. Of all observed movements into poverty, only 34 percent were coincident with the stoppage of labor income, and only 23 percent of the exits occurred at the same time as the initial receipt of Social Security benefits. While much of the movement within the income distribution does not appear to be attributable to a single cause, there is a large change in observed income associated with a transition into either state. Those entering poverty had an associated mean change in income between the two periods of $3,059 (in 1979 dollars). For those who transit out of poverty, this figure was $3,283. By comparison, the average change in income for those who remained not poor was $631, and those who remained poor experienced an average increase of $112. A detailed investigation of the factors causing movement into and out of poverty is an important topic, but beyond the scope of this paper.[5]

## IV. The Model

One aim of this paper is to obtain a correct estimate of the fraction of those in poverty at a given time who are expected to remain poor. Those individuals whose income is expected to remain below the poverty line indefi-

4. As the Social Security law is written, individuals may first collect a widow's benefit at age 60 (disabled widows may receive benefits as early as age 50). However, these payments are reduced relative to the amount which would be obtained if the receipt of benefits were begun at age 65.

5. Holden, Burkhauser, and Feaster (1988) include a limited set of explanatory variables in a hazard model of the probability of entering poverty. The effect of these characteristics on the likelihood of becoming poor provides some indication of the determinants of falls into poverty.

nitely will be termed the *chronically poor*. To examine the relative permanence of poverty, it is helpful to decompose the variation in income into random individual effects and stochastic elements. A variance components model provides an effective method for assessing the contribution of each of these factors to the total variance of income. The model used is based on those in Lillard and Willis (1978) and Lillard and Weiss (1979). It has the form

$$\ln(y_{it}) = \Gamma_t + \delta_i + \varepsilon_i \sqrt{t} + \upsilon_{it} + u_{it}$$

where $\ln(y_{it})$ is the natural log of observed income for individual $i$ at time $t$.

$\Gamma_t$ is a time effect which allows for changing economic conditions across time periods. The real value of Social Security benefits increased substantially during the 1970s. There were also both periods of inflation and recession. This term accounts for these and any other systematic year effects which were common across individuals.[6]

$\delta_i$ is a random individual specific effect. It represents individual differences in income levels which persist over time.

$\varepsilon_i$ is a second individual component which allows for differences in the way in which individual income changes over time. The interaction of this random individual effect with time, together with the random component, $\delta_i$, permits each individual to follow an income path with a unique slope and intercept.

Because factors which affect the individual level of income are also likely to influence the way in which income changes over time, these two components, $\delta_i$ and $\varepsilon_i$, are permitted to have a nonzero covariance. As mentioned in the previous section, it is theorized that those with exceptionally high income from earnings will retire at some point during the survey. Though income need not decrease with retirement, it would be unlikely to increase. Furthermore, it is expected that those who are poor will benefit disproportionately from government transfer programs, and experience a rise in real income. If these explanations are valid descriptions of the observed movements in the income distribution, then one would expect there to exist a negative correlation between $\delta_i$ and $\varepsilon_i$; those with low levels of income experience an increase over time, and those with high initial values tend to experience a decline. The two components are assumed to be distributed with zero means, and an unknown covariance matrix. The following notation is used:

$$\begin{pmatrix}\delta\\ \varepsilon\end{pmatrix} \sim \left(\begin{pmatrix}0\\ 0\end{pmatrix}, \begin{pmatrix}\sigma_\delta^2 & \sigma_{\delta\varepsilon}\\ \sigma_{\delta\varepsilon} & \sigma_\varepsilon^2\end{pmatrix}\right).$$

The time trend is modeled as a function of $\sqrt{t}$, rather than $t$ itself because this specification fits the data better than any reasonable alternative. The Appendix contains a discussion of the procedure used to evaluate the fit of alternative forms of the model.

$\upsilon_{it}$ captures the movements off the individual path implied by $\delta_i$ and $\varepsilon_i$, the individual specific slope and intercept. The sample covariances suggests the exis-

6. Modeling income as a function of age rather than time leads to almost identical results. These estimates will be discussed later.

tence of serial correlation (see appendix Table 1A). If there were no time trend or serially correlated disturbance, the covariance between the income reported at time $t$, $y_{it}$ and that reported at time $\tau$, $y_{i\tau}$ would be identical for all $t \neq \tau$. Rather than immediately falling to a fixed level, however, the covariance tends to decline over time. To model this feature of the data, $\upsilon_{it}$ is permitted to follow an AR(1) process.[7] Denoting the serial correlation coefficient by $\rho$, the autoregressive process is written as

$$\upsilon_{it} = \rho\upsilon_{it-1} + \eta_{it}$$

where

$$\eta_{it} \sim \text{iid}(0, \sigma^2_\eta).$$

Thus,

$$\sigma^2_\upsilon = \sigma^2_\eta/(1 - \rho^2).$$

The component $\upsilon_{it}$ can be thought of as accounting for shocks to income, the effects of which carry over to future periods, as well as for individual effects which change gradually over time.

Because observations in the RHS are two years apart, a higher order process did not seem warranted. An AR(2) process implies that a shock to income in one period will continue to influence income four years later. It is doubtful that any effect which is not captured by the individual components would be that persistent. Hence, an AR(1) process is expected to be sufficient.

The choice of the appropriate order autoregressive process does not follow from economic theory, but depends simply on which specification provides a better fit. MaCurdy (1982) suggests the use of formal time series techniques (inspection of the sample covariogram and the sample partial correlation function) to determine the appropriate autoregressive process. As is the case with typical time series models, the choice of the best fitting model is somewhat subjective. Rather than this formal procedure, I use a more straightforward, though less comprehensive approach, and estimate the model for several alternative specifications [including ARMA(1, 1) and AR(2) processes]. The coefficients for higher order processes were not significant at a 5 percent level. The estimates are included in an earlier version of this paper (1991).

$u_{it}$ is a purely stochastic, zero mean component which includes errors in the observed value of income. This specification is derived from the typical measurement error framework wherein the observed value is equal to the sum of the true value and a zero mean error term. The model can be thought of as

$$\ln(y_{it}) = \ln(y^*_{it}) + u_{it}$$

where

$$u_{it} \sim \text{iid}(0, \sigma^2_u).$$

7. Other work suggests that much of the serial correlation stems from employment earnings and/or the shift from employment to retirement. With a similarly selected sample of non-workers, the estimate of $\rho$ was insignificant (see McGarry 1992)).

The variable $y_{it}$ is observed income and $y^*_{it}$ represents true income; $y^*_{it}$ includes the year specific terms and all components other than $u_{it}$.

Although $u_{it}$ is depicted solely as measurement error, it also contains purely stochastic disturbances not captured by the other components of the variance. Thus, an estimate of $\sigma^2_u$ is actually an estimate of the sum of measurement error and the non-serially correlated stochastic component of income. The model cannot distinguish between measurement error and stochastic shocks to income which last for one period. However, if one is interested in a measure of the amount of permanence of poverty, not much is lost by grouping the two stochastic components of $u_{it}$. Regardless of the origin of the shock, be it measurement error or a true disturbance, its effect is temporary (unlike that of $\eta_{it}$) and does not alter the individual's expected future income.[8]

The model is identified because $\sigma_u$ enters into the calculation of the diagonal elements of the covariance matrix, but does not influence the off diagonal terms; $\sigma_\eta$ and $\rho$ appear in all elements. Ignoring the individual components for the moment, the covariance between the income at time $t$ and that at time $\tau$ can be written as

$$\text{cov}(y_{it}, y_{i\tau}) = \rho^{|t-\tau|}\sigma^2_\eta/(1 - \rho^2) + \sigma^2_u D_{it}$$

where

$$\begin{aligned} D_{it} &= 1 \text{ if } t = \tau \\ &= 0 \text{ elsewhere.} \end{aligned}$$

The first term of the covariance formula changes by a factor of $\rho$ each time the length of time between the two observations increases by one period. Thus $\rho$ is identified.[9]

Explanatory variables such as age, schooling, or experience, which are typically included in wage equations, do not appear in the model. This exclusion stems primarily from the focus of the paper. I am concerned with the extent to which individuals are likely to be faced with permanent as opposed to transitory spells of poverty, rather than with explaining the likelihood of entering poverty in terms of observed characteristics. Therefore, whether these individual differences are included as regressors or incorporated in the random individual component, the implications remain substantially unchanged. Furthermore, because the sample is comprised exclusively of widows, knowledge of the typical explanatory variables does not provide as much information on income as it does in models of earnings determination. Much of the income for these women comes from sources other than earnings. In many instances the amount of this income is determined to a significant extent by their late husband's income. Social Security payments, for example, are based on lifetime contributions. Because women in this cohort did not have a strong attachment to the labor force, the observed

8. Because the respondents in the RHS were surveyed only every other year, it is unlikely that response errors would be correlated over time.

9. In fact the model is over-identified. The identification can be verified by writing out the expression for covariance matrix and solving for the parameters in terms of the sample covariances.

payment is likely to be a function of their husbands' earnings. Similarly, pensions are usually determined by the occupation and earnings of the deceased husband.[10] Asset and property income also depend indirectly on the lifetime income of the couple. Hence, the financial status of these widows is in a large part determined by unobserved variables.[11]

A potentially interesting demographic variable which is excluded is race. The race of a widow is a good predictor of her husband's race and is, therefore, likely to be related to her financial status. Unfortunately, there are not enough non-whites in the sample to permit a study of racial differences in income dynamics. As a result race as well as other demographic characteristics is subsumed in the individual components $\delta_i$ and $\varepsilon_i$.

Given this specification of the model, the total variance can be divided into a portion due to permanent differences between individuals, a portion due to the serially correlated term, and a portion attributable to the purely transitory component which is assumed to include measurement error. The relative importance of each of these components is investigated in the next section. Estimates of the level of permanence of poverty and the likelihood of being observed in a sequence of poor/not poor states are also obtained.

## V. Estimation

Lillard and Willis (1978) assumes joint normality for all terms and estimates a model by maximum likelihood. However, they acknowledge that the distribution is not truly normal. The distribution of the log of income is typically found to be negatively skewed and leptokurtic. Gottschalk and Danziger (1985) use a three parameter log normal distribution to correct for the skewness in their sample. With this parameterization, they fail to reject the null hypothesis of normality, but only for a given range of the distribution. Bound and Krueger (1991), in a study examining the distribution of errors in reported income, find a large spike at the mean value of zero. They use a mixing distribution to estimate the model, and compare this result to that based on the assumption of normality. Their findings lead them to believe that the results are not greatly affected by the distributional assumption.

For this sample of widows from the RHS, the null hypothesis of normality is rejected at the one percent level for most years using both the Kolmogorov-Smirnov test and Lilliefors' test for normality [see Conover (1980) for an explanation of the latter procedure]. While it is clearly desirable to estimate the model without imposing distributional assumptions, my attempt at such estimation, us-

10. The relatively large number of women with reported earnings in 1969 does not imply high lifetime participation rates, and Social Security and pensions are based on lifetime contributions. Even for those women who did work at some point, it is likely that unless their husbands died at very young ages, Social Security benefits based on their own contributions would be less than those based on their husbands' earnings records.

11. Because these women were widowed before the survey began, their current financial status may be less dependent on that of their deceased husbands than would be the case for more recently widowed women.

ing a nonlinear least squares procedure, was unsuccessful. A Monte Carlo experiment revealed the source of the problem: because empirical identification becomes increasingly difficult as $\rho$ becomes small, estimates can only be obtained if the sample size is very large. Because my sample is not large enough (given the magnitude of the serial correlation coefficient), I am unable to estimate the model without the additional structure given by an assumed distribution. Therefore, I assume normality.[12]

To assess the impact of the normality assumption when the true data are nonnormal, I again employ a series of Monte Carlo experiments. I make various alternative distributional assumptions for the components of income, and generate several synthetic data sets by drawing from these distributions. The model is then estimated with the synthetic data using maximum likelihood. The likelihood function is that which would be appropriate if the distributions were truly normal. In all cases, the parameter estimates compare well with the known values, and the predictive abilities of the model do not appear to be sensitive to the distributions used to generate the data. I conclude that for the range of distributions and parameters investigated, the results obtained from the model are not greatly affected by the underlying distributions.[13]

### *Results*

Table 1 presents parameter estimates for the specification described above, as well as for one using dummy variables for each observed age. A final specification includes both age and year dummies. In neither case are the estimates significantly altered. When income is modeled in terms of age alone, the value of the likelihood function fails to improve; it increases only slightly with both year and age dummies. For each of the three models, the estimated values of the parameters are all significant at conventional levels. The estimated covariance between $\delta_i$ and $\varepsilon_i$ is negative, as expected. Because the estimates for each specification are almost identical, only the results for the first parameterization are used in the ensuing analysis.[14]

The portion of the variance attributable to the different components is reported in Table 2. The total variance in terms of the parameters is

$$\sigma_\delta^2 + t\sigma_\varepsilon^2 + 2\sqrt{t}\,\sigma_{\delta\varepsilon} + \sigma_\upsilon^2 + \sigma_u^2.$$

12. An example of an alternative estimation procedure can be found in Abowd and Card (1989). These authors use an optimal minimum distance estimator and an equally weighted distance estimator neither of which depends on distributional assumptions. Because the maximum likelihood method produces consistent estimates (albeit with incorrect standard errors), even if the distribution is not normal (MaCurdy 1982), the use of an alternative estimation procedure would not be expected to have a substantial impact.

13. The outcomes of the Monte Carlo experiments discussed in this and in the preceding paragraph are contained in an unpublished version of this paper (1991).

14. The chi-square statistics reject each of the three specifications. However, this result should not be interpreted to mean that the model is not useful in explaining the data. With sufficiently large samples, such tests are quite powerful and likely to reject the specification even when departures from the model are small. A better measure of the fit of the model is a comparison of observed and predicted poverty rates which will be discussed below.

**Table 1**
*Parameter Estimates for each Specification (standard errors in parentheses)*

| Parameter | Year Dummies | Age Dummies | Age and Year Dummies |
|---|---|---|---|
| $\sigma_\delta$ | 1.1132 | 1.1000 | 1.1034 |
| | (0.0455) | (0.0507) | (0.0447) |
| $\sigma_\varepsilon$ | 0.3400 | 0.3292 | 0.3351 |
| | (0.0321) | (0.0376) | (0.0315) |
| $\sigma_{\delta\varepsilon}$ | −0.3520 | −0.3400 | −0.3445 |
| | (0.0417) | (0.0462) | (0.0405) |
| $\sigma_\eta$ | 0.3342 | 0.3330 | 0.3270 |
| | (0.0495) | (0.0462) | (0.0429) |
| $\rho$ | 0.4554 | 0.4925 | 0.4759 |
| | (0.0262) | (0.1809) | (0.1543) |
| $\sigma_u$ | 0.3866 | 0.3908 | 0.3929 |
| | (0.0413) | (0.0383) | (0.0340) |
| Log likelihood | −5,281.6 | −5,282.7 | −5,274.8 |
| $\chi^2$ | 315 | 317 | 301 |

$\sigma_\delta$ is the standard deviation of the individual specific level component.
$\sigma_\varepsilon$ is the standard deviation of the individual specific time trend.
$\sigma_{\delta\varepsilon}$ is the covariance of the individual specific components.
$\sigma_\eta$ is the standard deviation of the innovation in the serially correlated term.
$\rho$ is the correlation coefficient.
$\sigma_u$ is the standard deviation of the white noise noise component.
$u$ represents estimated measurement error.

**Table 2**
*Contribution of Error Components to Total Variance*

| Year | Sample Variance | Estimated Variance | Percentage Due to Individual | Error | Other | Total | Reliability Ratio |
|---|---|---|---|---|---|---|---|
| 1969 | 0.945 | 0.941 | 69.2 | 15.9 | 15.0 | 100.0 | 0.841 |
| 1971 | 0.946 | 0.765 | 62.1 | 19.5 | 18.4 | 100.0 | 0.805 |
| 1973 | 0.692 | 0.657 | 55.8 | 22.8 | 21.5 | 100.0 | 0.773 |
| 1975 | 0.533 | 0.584 | 50.3 | 25.6 | 24.1 | 100.0 | 0.744 |
| 1977 | 0.492 | 0.533 | 45.5 | 28.0 | 26.4 | 100.0 | 0.720 |
| 1979 | 0.388 | 0.498 | 41.8 | 30.0 | 28.3 | 100.0 | 0.700 |

The sample variance and the estimated total variance are provided in the first two columns of Table 2. The sample variance remains relatively constant for the first two years and then declines sharply, most likely because of retirement. The structure of the model, however, implies a smoother change and does not capture the sudden drop.[15]

The total variance of income can be decomposed into the variance across individuals and the variance of an individual's income over time. The across-person variance is denoted as

$$\sigma_\delta^2 + t\sigma_\varepsilon^2 + 2\sqrt{t}\sigma_{\delta\varepsilon}.$$

The portion of the within-person variance due to the variation of true income is $\sigma_\upsilon^2$, the variance of the serially correlated component. In terms of the estimated parameters this value can be written as

$$\frac{\sigma_\eta^2}{(1 - \rho^2)}.$$

The remainder of the within-person variance is due to the stochastic component $u_{it}$. Its estimated value sets an upper bound on the portion of the variance attributable to measurement error.

Table 2 reveals that in the early years of the survey, a somewhat larger portion of the variance is due to individual differences than in the later years. Initially, 69 percent of the variance in income is accounted for by the individual effects. This value decreases to 42 percent by 1979. The declining importance of these components is in agreement with the notion that transfer programs and retirement act jointly to reduce the inequality in income. As the fraction of the variance accounted for by individual effects decreases, the remaining portion of the variance, divided fairly evenly between the serially correlated term and white noise ($\sigma_\upsilon^2$ and $\sigma_u^2$), increases.

The importance of spurious transitions in the number of observed exits from poverty depends on the size of the observed error relative to the reported change in income. The ratio of the variance of true income (excluding measurement error) to the variance of observed income is the *reliability ratio*. If the random white noise component ($u_{it}$) is assumed to represent the error in observed income, the reliability ratio in this model can be written as

$$\frac{\sigma_\delta^2 + \sigma_\upsilon^2 + t\sigma_\varepsilon + 2\sqrt{t}\,\sigma_{\delta\varepsilon}}{\sigma_\delta^2 + \sigma_\upsilon^2 + t\sigma_\varepsilon + 2\sqrt{t}\,\sigma_{\delta\varepsilon} + \sigma_u^2}.$$

15. I also estimated the model without the individual specific time trend, but with an additional parameter in each year to allow the variance of income to change randomly across years. I prefer the model presented above because it implies that each individual follows an income path with a unique slope and intercept, in accord with the earnings literature. It also allows for the calculation of statistics such as the proportion who are chronically poor.

The values for each period are reported in the rightmost column of Table 2.[16] Since $\sigma_u^2$ is constant over time and the sample variance is decreasing, the fraction of the total variance attributable to this component increases, and the reliability decreases. Initially the reliability is 0.84, but by 1979 this ratio decreases to 0.70.

## VI. Implications for Poverty Rates

In each period, the fraction of the population that is expected to have income below the poverty line is

$$P(\Gamma_t + \delta_i + \varepsilon_i \sqrt{t} + \upsilon_{it} + u_{it} < \ln(POV))$$

where $POV$ is the poverty line, \$3,479 in 1979.

Under the assumption of joint normality, this probability can be written as

$$(1) \quad \Phi\left(\frac{\ln(POV) - \Gamma_t}{\sqrt{\sigma_t^2}}\right)$$

where $\Phi$ denotes the cumulative distribution function and

$$\sigma_t^2 = \sigma_\delta^2 + t\sigma_\varepsilon^2 + 2\sqrt{t}\,\sigma_{\delta\varepsilon} + \sigma_\upsilon^2 + \sigma_u^2.$$

Table 3 shows the observed and the predicted proportions in poverty for each year. Some of the discrepancy between the two values stems from imperfect estimation of the sample variance, and some from the distributional assumption. To obtain an indication of the size of the error associated with the distributional assumption, I calculate the expected poverty rate using (1), but with the sample mean and variance rather than the estimates obtained from the model. The difference between the observed proportion in poverty and these predicted values is due solely to the distributional assumption. The size of the error is on the order of 2 percentage points. The difference results because there are fewer widows with incomes in the lower tail of the distribution than is implied by the normal distribution. The predicted poverty rate is higher than the observed rate in all but the last year. The same table shows the predicted proportions calculated with

16. These ratios are comparable to those found elsewhere. The Census estimate of the reliability ratio for income is approximately 0.85 (Fuller 1987). Bound and Krueger (1991) calculate a reliability ratio for earnings of about 0.82 for men and 0.92 for women. Duncan and Hill (1985) examine earnings and estimate reliabilities of between 0.85 and 0.68. Since the error term here is comprised both of measurement error and genuine stochastic variation, the reliability would be expected to be somewhat less than that found in other studies. However for a number of reasons, the true reliability for this sample might be somewhat larger. First, as demonstrated by Bound and Krueger, women seem to report more accurately than men. Also, while the Census permits proxy responses, all of the data for the widows in this sample are self-reported. One would expect self-reported measures of income to be subject to less response error than income measures reported by other family members. Furthermore, a substantial fraction of an elderly person's income is derived from stable sources such as Social Security and pensions. Thus, income in this sample ought to be less volatile than earnings, and perhaps more accurately reported. The net effect of lower response error and an estimate of $\sigma_u^2$ which includes both measurement error and other stochastic variation is indeterminate. However, the estimated values seem entirely plausible.

**Table 3**
*Alternative Estimates of the Proportion Poor*

| Year | Actual Sample Proportion | Estimated with Sample Statistics | Estimated from Model: $u_{it}$ Included[a] | Estimated from Model: $u_{it}$ Excluded | Proportion Chronically Poor | Chronically Poor as a Fraction of All Poor |
|---|---|---|---|---|---|---|
| 1969 | 0.312 | 0.348 | 0.348 | 0.335 | 0.319 | 0.917 |
| 1971 | 0.333 | 0.356 | 0.340 | 0.323 | 0.301 | 0.883 |
| 1973 | 0.305 | 0.329 | 0.324 | 0.302 | 0.271 | 0.835 |
| 1975 | 0.318 | 0.328 | 0.335 | 0.311 | 0.274 | 0.818 |
| 1977 | 0.337 | 0.345 | 0.351 | 0.326 | 0.285 | 0.812 |
| 1979 | 0.397 | 0.355 | 0.371 | 0.347 | 0.306 | 0.823 |

a. $u_{it}$ represents estimated measurement error.

the estimated parameters. The model predicts poverty rates close to the observed percentages, although it again appears that poverty is over-estimated in all but the last year. The magnitude of the difference is similar to that obtained using the sample mean and variance.

To assess the effect of white noise errors on these estimates, I exclude the stochastic variation in income by setting the variance of this component, $\sigma_u^2$, equal to zero. Because true random shocks to income are also excluded by this procedure, the correct poverty rates lie between those obtained with $\sigma_u^2$ equal to zero and those calculated with the estimated value of $\sigma_u^2$. This procedure, therefore, provides a lower bound for the true poverty rates. Table 3 shows that eliminating the portion of the variance attributable to measurement error decreases the estimated proportion in poverty by less than 2 percentage points, on average. This result is not surprising. Because the poverty line lies in the left tail of the distribution, decreasing the variance of income decreases the proportion of the distribution below the poverty line. This small change emphasizes the fact that cross sectional estimates of the poverty rate are not seriously biased by the existence of measurement error.

Using a term from Bane and Ellwood (1986), I define the *chronically poor* as those whose expected incomes are below the poverty line. With *POV* denoting the poverty line, the chronically poor are those whose draws on the individual components, $\delta_i$ and $\varepsilon_i$ are such that

$$\Gamma_t + \delta_i + \varepsilon_i \sqrt{t} < POV.$$

The proportion of the population satisfying this inequality in year $t$ is

$$P(\Gamma_t + \delta_i + \varepsilon_i \sqrt{t} < POV).$$

If the joint distribution of $\delta_i$ and $\varepsilon_i$ is bivariate normal, this fraction is

$$\Phi\left(\frac{POV - \Gamma_t}{\sqrt{\sigma_{t*}^2}}\right)$$

where

$$\sigma_{t^*}^2 = \sigma_\delta^2 + t\sigma_\varepsilon^2 + 2\sqrt{t}\sigma_{\delta\varepsilon}.$$

The rightmost columns of Table 3 present the proportion of the population expected to be chronically poor and the fraction of those below the poverty line who belong to this group. The model predicts that well over 80 percent of those in poverty in a given year will remain there indefinitely. This statement is not equivalent to saying that 80 percent of those entering poverty in a single year will remain there permanently, but rather that the population below the poverty line at any time is composed substantially of the chronically poor. Recall that the estimate of the portion of the variance attributable to permanent differences among individuals is decreasing over time (see Table 2 and the previous discussion). Because of the shape of the normal distribution, and the location of the poverty line, the decrease in the individual specific variance implies that the fraction of those below the poverty line who are expected to remain poor indefinitely also decreases.

Table 4 shows the number of women who are observed in poverty for a given number of years. Just over 12 percent of the women in the sample are poor in all six years. An even larger portion experiences only marginally less adverse circumstances. While not spending every year in poverty, 27.2 percent are poor in at least four of the six years, and only 39.4 percent are never observed in poverty.

Because the number of years that a widow is observed in poverty is likely to be affected by measurement error, the true distribution may be very different. The distribution of years in poverty predicted by the model is also shown in Table 4.[17] Estimates are presented with the measurement error component both included and excluded in the calculation of income. When this approximation of measurement error is eliminated, the percentage of women experiencing poverty in all six years increases substantially, from 12.4 percent to 19.2 percent. The percent not poor in all years increases by an even larger amount, from 19.9 to 28.4 percent. The difference in the effect of measurement error on the number of "always poor" and "never poor" follows from the previous discussion of its impact on poverty rates. Setting $\sigma_u^2$ equal to zero reduces the proportion in poverty at any one time and, therefore, the proportion which remains poor in all periods. In addition to this change, there also is a reduction in the number of spurious transitions both into and out of poverty. When examining the permanently poor, the two effects work in opposite directions. However, with respect to those who are never poor, these effects work together and the change is more pronounced.

Table 4 emphasizes an advantage of this variance components method over a procedure which discounts one period spells of poverty or non-poverty. The variance components model predicts that 16.3 percent of the sample will spend

17. To calculate these values, I use the estimated parameters of the income distribution to generate synthetic data. From this data, the expected number of women spending any number of years in poverty can be observed.

**Table 4**
*Number of Years Spent in Poverty*

| | | | Predictions from Simulations | | | |
|---|---|---|---|---|---|---|
| | From Actual Data | | $u_{it}$ Included[a] | | $u_{it}$ Excluded | |
| Number of Years | Number | Percent | Number | Percent | Number | Percent |
| 0 | 368 | 39.4 | 186 | 19.9 | 265 | 28.4 |
| 1 | 154 | 16.5 | 152 | 16.3 | 121 | 13.0 |
| 2 | 75 | 8.0 | 125 | 13.4 | 90 | 9.6 |
| 3 | 83 | 8.9 | 101 | 10.8 | 83 | 8.9 |
| 4 | 66 | 7.1 | 123 | 13.2 | 101 | 10.8 |
| 5 | 75 | 8.0 | 131 | 14.0 | 95 | 10.2 |
| 6 | 113 | 12.1 | 116 | 12.4 | 179 | 19.2 |
| Total | 934 | 100 | 934 | 100 | 934 | 100 |

a. $u_{it}$ represents estimated measurement error.

just one year in poverty. This figure compares well with the 16.5 percent of the sample who are actually observed in poverty for a single period.[18] Within the framework of the variance components model, one period spells are observed even after a correction for estimated measurement error. When $\sigma_u^2$ equals zero, approximately 13 percent of the widows in the sample are expected to experience a single year of poverty, and 10 percent are expected to have income above the poverty line for only one year. While these percentages are reduced relative to their values before the correction, one period spells continue to categorize the poverty experience of a large portion of the sample. Eliminating these transitions conceals an important feature of the data.

Tables 3 and 4 show that a large number of individuals are expected to be faced with many years of poverty. However, it is also apparent that many widows experience short, infrequent spells of poverty. Thus, while most of those who become poor in any period will exit shortly, the majority of those in poverty at a given time will be poor permanently. In each period this group of permanently poor is joined by a small number of the transient poor.

The conditional probabilities of observing a widow in poverty in a particular year, given that she was poor in any preceding year, are presented in Table 5. The observed sample values are presented first, followed by the values predicted by the model. The conditional probabilities are calculated for both the full model and the model with estimated measurement error excluded. When measurement error is eliminated, the conditional probability of being poor in one year given

18. Because interviews in the RHS are conducted every other year rather than annually, it is possible that a widow who is observed in poverty in only one of the six surveys, actually spends some of the unobserved time in poverty as well.

**Table 5**
*Estimated Conditional Probabilities Probability of Being Poor in Year Y Conditional on Being Poor in Year X*

| | Year $Y$ | | | | |
|---|---|---|---|---|---|
| Year $X$ | 1971 | 1973 | 1975 | 1977 | 1979 |
| 1969 | | | | | |
| Observed | 0.749 | 0.615 | 0.608 | 0.653 | 0.701 |
| Model including $u_{it}^{a}$ | 0.671 | 0.604 | 0.582 | 0.570 | 0.565 |
| Model excluding $u_{it}$ | 0.773 | 0.676 | 0.643 | 0.622 | 0.611 |
| 1971 | | | | | |
| Observed | | 0.656 | 0.653 | 0.691 | 0.740 |
| Model including $u_{it}$ | | 0.625 | 0.592 | 0.577 | 0.572 |
| Model excluding $u_{it}$ | | 0.730 | 0.672 | 0.644 | 0.631 |
| 1973 | | | | | |
| Observed | | | 0.677 | 0.702 | 0.750 |
| Model including $u_{it}$ | | | 0.623 | 0.593 | 0.584 |
| Model excluding $u_{it}$ | | | 0.742 | 0.683 | 0.660 |
| 1975 | | | | | |
| Observed | | | | 0.795 | 0.805 |
| Model including $u_{it}$ | | | | 0.614 | 0.591 |
| Model excluding $u_{it}$ | | | | 0.734 | 0.682 |
| 1977 | | | | | |
| Observed | | | | | 0.835 |
| Model including $u_{it}$ | | | | | 0.613 |
| Model excluding $u_{it}$ | | | | | 0.734 |

a. $u_{it}$ represents estimated measurement error.

that one is poor in the previous period is increased by approximately 9 percentage points, or 13 percent. For the years 1969 and 1971, the increase is from 0.671 to 0.773. The observed probability is 0.749. In the sample, the number of those in poverty increases sharply in the last period. The model does not capture this large increase (see Table 2), so it also fails to capture the large increase in the conditional probability of being poor in 1979 given that one is poor in an earlier year. For the conditional probabilities in earlier years, the actual and predicted values also differ by a large amount. The true conditional probabilities, exclusive of measurement error would be expected to be approximately 13 percent greater than what is observed. Thus, the corrected probability of being poor in 1971 given that one is poor in 1969 would be expected to increase from 75 to 85 percent.

Approximately one-third of the widows are in poverty in any year, regardless of their poverty status in a prior period. However, conditional on being poor in the previous year, the proportion of women expected to be poor in the current

**Table 6**
*Estimated Joint Probabilities Probability of Being Poor in Year Y and Year X*

| | Year $Y$ | | | | |
|---|---|---|---|---|---|
| Year $X$ | 1971 | 1973 | 1975 | 1977 | 1979 |
| 1969 | | | | | |
| Observed | 0.233 | 0.192 | 0.190 | 0.203 | 0.218 |
| Model including $u_{it}^{a}$ | 0.233 | 0.210 | 0.202 | 0.198 | 0.196 |
| Model excluding $u_{it}$ | 0.259 | 0.226 | 0.215 | 0.208 | 0.205 |
| 1971 | | | | | |
| Observed | | 0.218 | 0.217 | 0.230 | 0.246 |
| Model including $u_{it}$ | | 0.213 | 0.201 | 0.196 | 0.194 |
| Model excluding $u_{it}$ | | 0.236 | 0.217 | 0.208 | 0.204 |
| 1973 | | | | | |
| Observed | | | 0.207 | 0.214 | 0.229 |
| Model including $u_{it}$ | | | 0.202 | 0.192 | 0.189 |
| Model excluding $u_{it}$ | | | 0.224 | 0.206 | 0.199 |
| 1975 | | | | | |
| Observed | | | | 0.253 | 0.256 |
| Model including $u_{it}$ | | | | 0.206 | 0.198 |
| Model excluding $u_{it}$ | | | | 0.228 | 0.212 |
| 1977 | | | | | |
| Observed | | | | | 0.282 |
| Model including $u_{it}$ | | | | | 0.215 |
| Model excluding $u_{it}$ | | | | | 0.239 |

a. $u_{it}$ represents estimated measurement error.

period is over 0.60. The exact values in each period correspond to the diagonal elements in Table 5. Knowledge of current poverty status is a strong predictor of the future probability of being poor.

The joint probability of being observed in poverty for any two years is reported in Table 6. Again, both the sample probabilities and the calculations based on the estimated model are presented. When $\sigma_u^2$ is included in the model, the predicted probability of being poor in two adjacent years is quite close to the observed value. Focusing again on the poverty rate for 1969 and 1971, the expected probability of being poor in both years is 0.233. This probability is identical to the observed probability. When $\sigma_u^2$ is excluded, the joint probability of being poor in two consecutive periods increases. For 1969 and 1971, for example, the probability increases by over 2 percentage points to 0.259.

A further indication of the amount of movement into and out of poverty can be obtained by examining the average transition probabilities. Table 7 shows the observed and predicted probabilities of an individual's poverty status given her poverty status in the previous period, averaged over all years. In each period the

**Table 7**
*Average Transition Probabilities*

| Status in Period $t$ | Status in Period $t + 1$: Not Poor | Poor |
|---|---|---|
| Not Poor | | |
| Observed | 0.8533 | 0.1467 |
| Model including $u_{it}^{a}$ | 0.8023 | 0.1977 |
| Model excluding $u_{it}$ | 0.8758 | 0.1242 |
| Poor | | |
| Observed | 0.2577 | 0.7423 |
| Model including $u_{it}$ | 0.3706 | 0.6294 |
| Model excluding $u_{it}$ | 0.2575 | 0.7425 |

a. $u_{it}$ represents estimated measurement error.

observed probability of leaving poverty is high. On average, approximately 25 percent of those in poverty in one year apparently exit by the following period. Measurement error has an important effect on these transition probabilities. When the measurement error term is included in the model, the average probability of entering poverty is 0.198, and the probability of exiting is 0.376. When the effect of this transitory component is eliminated, these values fall to 0.124, and 0.258. Although the proportion in poverty at a given time is changed only slightly by excluding measurement error, the change in the transition probabilities is substantial. As is the case with the conditional probabilities of Table 5, the model (with the measurement error term included) predicts transition probabilities which do not closely agree with the observed probabilities. To make the correction for measurement error to the sample transition matrices one would adjust the probabilities by the percentage change observed between the two versions of the model. Thus the probability of entering poverty would be reduced by 37 percent from approximately 0.15 to 0.095 and the probability of exiting poverty would fall from 0.74 to approximately 0.5.

## VII. An Alternative Correction Procedure

The correction for measurement error explored in this paper is not the only possible solution. A straightforward alternative is the method proposed by both Bound et al. and by Bane and Ellwood, wherein one period spells of poverty or non-poverty are treated as resulting from errors in reported income. Although the timing of the interviews in the RHS makes it difficult to replicate this alternative method with my sample, some type of comparison is warranted.

Because observations in the RHS are collected biennially, reports of yearly income, as used by Bound et al. are not available. I therefore compared the two

**Table 8**
*Average Transition Probabilities Synthetic Data*

| | Status in Period $t + 1$ | |
|---|---|---|
| Status in Period $t$ | Not Poor | Poor |
| Not Poor | | |
| Observed | 0.80 | 0.20 |
| Model including $u_{it}^{a}$ | 0.80 | 0.20 |
| Model excluding $u_{it}$ | 0.88 | 0.12 |
| Eliminating single spells | 0.94 | 0.06 |
| Poor | | |
| Observed | 0.34 | 0.66 |
| Model including $u_{it}$ | 0.37 | 0.63 |
| Model excluding $u_{it}$ | 0.25 | 0.75 |
| Eliminating single spells | 0.08 | 0.92 |

a. $u_{it}$ represents estimated measurement error.

methodologies using synthetic data. I used the estimated model to generate a series of 11 values of income for the years 1969–1979, for 934 hypothetical individuals. The constants for the non-survey years were obtained through interpolation (for example, $\Gamma_{70} = (\Gamma_{69} + \Gamma_{71})/2$). With this new data set I replicated the procedure in Bound et al., and derived the corresponding transition matrices and poverty rates, and I calculated the predicted probabilities generated by the model both with and without the measurement error component. The comparison to be made is between the two sets of corrected probabilities: those predicted by the model excluding the variance of the white noise component, and those obtained by eliminating one-period transitions.

Neither of these correction procedures has much effect on the cross sectional estimates of the proportion poor. Each reduces the proportion of those with incomes below the poverty line by approximately 2 percentage points. However with respect to the transitional probabilities there is a substantial difference. Four different calculations of these probabilities are presented in Table 8. The observed probability of becoming poor in one year, given that one is poor in the preceding year, averaged over all years, is approximately 20 percent. The observed probability of exiting poverty is 34 percent. The model predicts probabilities of 20 percent and 37 percent (these are the same values as are reported in Table 7). Because these data are generated from the model, these calculated probabilities are expected to be in close agreement with the observed probabilities. When the measurement error component is excluded from the determination of the variance, the model predicts transition probabilities of 12 and 25 percent. With the synthetic data and the assumption that all white noise disturbances are due to measurement error, these are the true transition probabilities.

Following the procedure used in Bound et al., the corrected transition probabilities are 6 and 8 percent. This comparison exaggerates the value of the model in that the model does not describe true income as accurately as it describes the synthetic values. With that limitation, however, I conclude that the elimination of all one-period spells leads to a substantial underestimate of the transition probabilities.

## VIII. Conclusion

Few economic variables provide as direct a link to an individual's well-being as measures of income. The poverty line provides a benchmark against which to compare this income. It provides a convenient means of comparing financial status across families of varying sizes, and of examining changes in the distribution of income over time. Accurate measurement of poverty rates is of further importance as published poverty statistics are often used to determine government policy. However, as the numbers presented above indicate, much of the observed movement about the poverty line may be caused by errors in the measurement of income. The number of spurious transitions into and out of poverty appears to be substantial. When transitory shocks are eliminated, the probability of exiting poverty decreases by just over 30 percent relative to the predictions of the full model. The probability of entering poverty decreases by about 37 percent. Despite the large reduction in the probability of entering or exiting poverty the amount of movement about the poverty line is still substantial: over 25 percent of those poor in a given period are expected to exit by the following period. It should again be noted, however, that my sample consists only of women who are between the ages of 58 and 63 in 1969 and who are already widowed at this point. Poverty measures for other samples would be expected to differ.

However, by examining measures of permanence, it is apparent that all widows do not face an equal probability of escaping from poverty. While the majority of those who experience a spell of poverty are poor only temporarily, most of those who are poor at any one time are in the midst of a long spell of poverty. When the estimated measurement error is eliminated from the model, the percent of women who are classified as poor in all six sample periods increases from 12.4 to 19.2 percent, while the percent never-poor increases from 19.9 to 28.4 percent.

The changes observed when the white noise component is eliminated demonstrate the importance of controlling for the existence of measurement error in studies of income dynamics. Although the amount of movement about the poverty line is large, even when transitory shocks to income are eliminated, the influence of the measurement error component appears to be substantial.

## Appendix

### *Alternative Specification*

An alternative specification of the model would formulate the time trend in terms of $t$. However, using successive integer values for $t$ implies that the income path is linear in $t$. The data do not appear to exhibit this feature. An obvious solution is to include $t^2$ in addition to $t$ to parameterize this non-linearity. However, if

**Table 1A**
*Correlation of Income Across Years*

| | 1969 | 1971 | 1973 | 1975 | 1977 | 1979 |
|---|---|---|---|---|---|---|
| 1969 | 1.000 | 0.661 | 0.578 | 0.493 | 0.468 | 0.514 |
| 1971 | 0.661 | 1.000 | 0.636 | 0.536 | 0.530 | 0.552 |
| 1973 | 0.578 | 0.636 | 1.000 | 0.634 | 0.577 | 0.596 |
| 1975 | 0.493 | 0.536 | 0.634 | 1.000 | 0.653 | 0.649 |
| 1977 | 0.468 | 0.530 | 0.577 | 0.653 | 1.000 | 0.686 |
| 1979 | 0.514 | 0.552 | 0.596 | 0.649 | 0.686 | 1.000 |

$t^2$ is included in the model, three additional parameters are required: one for the variance of the coefficient on $t^2$ and two more to handle the covariances between this term and the other individual effects ($\delta_i$ and $\varepsilon_i$). To capture the nonlinearities of the data in a more parsimonious way, I model the time trend as $\varepsilon_i \sqrt{t}$, where $t = 1, 2 \ldots 6$. To verify that $\sqrt{t}$ provides a better fit than $t$ alone, I examine the relative size of the bias introduced by each specification. I estimate the bias using the procedure described in Murphy and Welch (1990). Ordinary least squares regressions are run for the following three specifications:

$$\log(y_{it}) = \alpha_1 + \beta_{11} t + \beta_{12} t^2 + e_1$$

$$\log(y_{it}) = \alpha_2 + \beta_{21} t + e_2$$

$$\log(y_{it}) = \alpha_3 + \beta_{31} \sqrt{t} + e_3 .$$

Each equation is then weighted by $1/\hat{\sigma}_{e_j}$, where $\hat{\sigma}_{e_j}$ is the estimated standard deviation of the residual $e_j, j = 1, 2, 3$. Reestimating the equations using weighted least squares, $E[\Sigma\, e_j^{*2}] = N - K$. The expected mean squared error is therefore $\frac{N-K}{N}$. The estimated value differs from the expected value because of the bias introduced by the specification. The relative size of the bias is calculated as the difference between the estimated mean squared error and its expected value. In the first equation, the specification bias is calculated to be 85 percent of the total mean squared residual. When only $t$ is used this percentage jumps to 91 percent. The specification with $\sqrt{t}$ performs better than either of the other two models. The proportion of the error variance due to the bias is 83 percent. Although the difference between the alternative specifications is small, it seems sensible to choose the specification with the smallest relative bias. Therefore, because $\sqrt{t}$ appears to fit the data better than the alternatives, this specification is selected for the full model.

## References

Abowd, John M., and David Card. 1989. "On the Covariance Structure of Earnings and Hours Changes." *Econometrica* 57(1):411–45.

Bane, Mary Jo, and David T. Ellwood. 1986. "Slipping Into and Out of Poverty: The Dynamics of Spells." *Journal of Human Resources* 21(1):1–23.

Bound, John, and Alan B. Krueger. 1991. "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics* 9(1):1–24.

Bound, John, Greg J. Duncan, Deborah S. Laren, and Lewis Oleinick. 1991. "Poverty Dynamics in Widowhood." *Journal of Gerontology* 46(3):S115–S124.

Burkhauser, Richard, and Greg J. Duncan. 1990. "Public Policy and the Income Security of the Elderly." Paper presented at the ISPE Conference on the Fiscal Implications of an Aging Population, Limburg, The Netherlands.

Burkhauser, Richard V., Karen C. Holden, and Daniel A. Myers. 1986. "Marital Disruption and Poverty: The Role of Survey Procedures in Artificially Creating Poverty." *Demography* 23(4):621–31.

Coe, Richard D. 1978. "Dependency and Poverty in the Short and Long Run." In *Five Thousand American Families: Patterns of Economic Progress*, Vol. 6, ed. Greg J. Duncan and James N. Morgan, 273–96. Ann Arbor: Institute for Social Research, University of Michigan.

Conover, W. J. 1980. *Practical Nonparametric Statistics*. New York: John Wiley & Sons Inc.

Duncan, Greg J. 1984. *Years of Poverty, Years of Plenty: The Changing Economic Fortunes of American Workers*. Ann Arbor: University of Michigan Press.

Duncan, Greg J., and Daniel H. Hill. 1985. "An Investigation of the Extent and Consequences of Measurement Error in Labor-Economic Survey Data." *Journal of Labor Economics* 3(4):508–32.

Goldin, Claudia. 1990. *Understanding the Gender Gap*. New York: Oxford University Press.

Gottschalk, Peter, and Sheldon Danziger. 1985. "A Framework for Evaluating the Effects of Economic Growth and Transfers on Poverty." *The American Economic Review* 75(1):153–61.

Holden, Karen C., Richard V. Burkhauser, and Daniel J. Feaster. 1988. "The Timing of Falls into Poverty After Retirement and Widowhood." *Demography* 25(3):405–14.

Holden, Karen C., Richard V. Burkhauser, and Daniel A. Meyers. 1986. "Income Transitions at Older Stages of Life: The Dynamics of Poverty." *The Gerontologist* 26(3):292–97.

Hurd, Michael D., and David A. Wise. 1989. "The Wealth and Poverty of Widows: Assets Before and After the Husband's Death." In *The Economics of Aging*, ed. David A. Wise, 177–99. Chicago: NBER and the University of Chicago Press.

Lillard, Lee A., and Yoram Weiss. 1979. "Components of Variation in Panel Earnings Data: American Scientists 1960–1970." *Econometrica* 47(2):437–54.

Lillard, Lee A., and Robert J. Willis. 1978. "Dynamic Aspects of Earning Mobility." *Econometrica* 46(5):985–1010.

MaCurdy, Thomas E. 1982. The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis." *Journal of Econometrics* 18(1):83–114.

McGarry, Kathleen. 1992. "Changes in Income and Poverty Status in SIPP." Unpublished paper, UCLA.

———. 1991. "Measurement Error and Poverty Rates of Widows." Unpublished paper, UCLA.

Murphy, Kevin M., and Finis Welch. 1990. "Empirical Age-Earnings Profiles." *Journal of Labor Economics* 8(2):202–29.

Social Security Administration, U.S. Department of Health and Human Services. 1988. *Income of the Population Aged 55 and Over*. Washington, D.C.: GPO.

U.S. Bureau of the Census. 1988. "Marital Status and Living Arrangement: March 1987." *Current Population Reports, Consumer Income,* Series P-20, No. 423.

———. 1988. "Poverty in the United States 1987." *Current Population Reports, Consumer Income*, Series P-60, No. 163.

# LINKED CITATIONS



*You have printed the following article:*

**Measurement Error and Poverty Rates of Widows**
Kathleen McGarry
*The Journal of Human Resources*, Vol. 30, No. 1. (Winter, 1995), pp. 113-134.
Stable URL:
http://links.jstor.org/sici?sici=0022-166X%28199524%2930%3A1%3C113%3AMEAPRO%3E2.0.CO%3B2-0

---



## [Footnotes]

[3] **Marital Disruption and Poverty: The Role of Survey Procedures in Artificially Creating Poverty**
Richard Burkhauser; Karen C. Holden; Daniel A. Myers
*Demography*, Vol. 23, No. 4. (Nov., 1986), pp. 621-631.
Stable URL:
http://links.jstor.org/sici?sici=0070-3370%28198611%2923%3A4%3C621%3AMDAPTR%3E2.0.CO%3B2-J

[5] **The Timing of Falls Into Poverty After Retirement and Widowhood**
Karen C. Holden; Richard V. Burkhauser; Daniel J. Feaster
*Demography*, Vol. 25, No. 3. (Aug., 1988), pp. 405-414.
Stable URL:
http://links.jstor.org/sici?sici=0070-3370%28198808%2925%3A3%3C405%3ATTOFIP%3E2.0.CO%3B2-B

[12] **On the Covariance Structure of Earnings and Hours Changes**
John M. Abowd; David Card
*Econometrica*, Vol. 57, No. 2. (Mar., 1989), pp. 411-445.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198903%2957%3A2%3C411%3AOTCSOE%3E2.0.CO%3B2-U

**NOTE:** *The reference numbering from the original has been maintained in this citation list.*

# LINKED CITATIONS



[16] **The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?**
John Bound; Alan B. Krueger
*Journal of Labor Economics*, Vol. 9, No. 1. (Jan., 1991), pp. 1-24.
Stable URL:
http://links.jstor.org/sici?sici=0734-306X%28199101%299%3A1%3C1%3ATEOMEI%3E2.0.CO%3B2-T

[16] **An Investigation of the Extent and Consequences of Measurement Error in Labor-Economic Survey Data**
Greg J. Duncan; Daniel H. Hill
*Journal of Labor Economics*, Vol. 3, No. 4. (Oct., 1985), pp. 508-532.
Stable URL:
http://links.jstor.org/sici?sici=0734-306X%28198510%293%3A4%3C508%3AAIOTEA%3E2.0.CO%3B2-R

## References

**On the Covariance Structure of Earnings and Hours Changes**
John M. Abowd; David Card
*Econometrica*, Vol. 57, No. 2. (Mar., 1989), pp. 411-445.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198903%2957%3A2%3C411%3AOTCSOE%3E2.0.CO%3B2-U

**Slipping into and out of Poverty: The Dynamics of Spells**
Mary Jo Bane; David T. Ellwood
*The Journal of Human Resources*, Vol. 21, No. 1. (Winter, 1986), pp. 1-23.
Stable URL:
http://links.jstor.org/sici?sici=0022-166X%28198624%2921%3A1%3C1%3ASIAOOP%3E2.0.CO%3B2-S

**The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?**
John Bound; Alan B. Krueger
*Journal of Labor Economics*, Vol. 9, No. 1. (Jan., 1991), pp. 1-24.
Stable URL:
http://links.jstor.org/sici?sici=0734-306X%28199101%299%3A1%3C1%3ATEOMEI%3E2.0.CO%3B2-T

**NOTE:** *The reference numbering from the original has been maintained in this citation list.*

# LINKED CITATIONS



**Marital Disruption and Poverty: The Role of Survey Procedures in Artificially Creating Poverty**
Richard Burkhauser; Karen C. Holden; Daniel A. Myers
*Demography*, Vol. 23, No. 4. (Nov., 1986), pp. 621-631.
Stable URL:
http://links.jstor.org/sici?sici=0070-3370%28198611%2923%3A4%3C621%3AMDAPTR%3E2.0.CO%3B2-J

**An Investigation of the Extent and Consequences of Measurement Error in Labor-Economic Survey Data**
Greg J. Duncan; Daniel H. Hill
*Journal of Labor Economics*, Vol. 3, No. 4. (Oct., 1985), pp. 508-532.
Stable URL:
http://links.jstor.org/sici?sici=0734-306X%28198510%293%3A4%3C508%3AAIOTEA%3E2.0.CO%3B2-R

**A Framework for Evaluating the Effects of Economic Growth and Transfers on Poverty**
Peter Gottschalk; Sheldon Danziger
*The American Economic Review*, Vol. 75, No. 1. (Mar., 1985), pp. 153-161.
Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28198503%2975%3A1%3C153%3AAFFETE%3E2.0.CO%3B2-F

**The Timing of Falls Into Poverty After Retirement and Widowhood**
Karen C. Holden; Richard V. Burkhauser; Daniel J. Feaster
*Demography*, Vol. 25, No. 3. (Aug., 1988), pp. 405-414.
Stable URL:
http://links.jstor.org/sici?sici=0070-3370%28198808%2925%3A3%3C405%3ATTOFIP%3E2.0.CO%3B2-B

**Components of Variation in Panel Earnings Data: American Scientists 1960-70**
Lee A. Lillard; Yoram Weiss
*Econometrica*, Vol. 47, No. 2. (Mar., 1979), pp. 437-454.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28197903%2947%3A2%3C437%3ACOVIPE%3E2.0.CO%3B2-M

**Dynamic Aspects of Earning Mobility**
Lee A. Lillard; Robert J. Willis
*Econometrica*, Vol. 46, No. 5. (Sep., 1978), pp. 985-1012.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28197809%2946%3A5%3C985%3ADAOEM%3E2.0.CO%3B2-B

**NOTE:** *The reference numbering from the original has been maintained in this citation list.*

**Empirical Age-Earnings Profiles**
Kevin M. Murphy; Finis Welch
*Journal of Labor Economics*, Vol. 8, No. 2. (Apr., 1990), pp. 202-229.
Stable URL:
http://links.jstor.org/sici?sici=0734-306X%28199004%298%3A2%3C202%3AEAP%3E2.0.CO%3B2-5