Climbing Out of Poverty, Falling Back In

Measuring the Persistence of Poverty Over Multiple Spells

Ann Huff Stevens

ABSTRACT

This paper investigates the persistence of poverty over individuals' lifetimes using a hazard rate approach that accounts for multiple spells of poverty and incorporates spell duration, individual and household characteristics, and unobserved heterogeneity. The findings highlight the importance of considering multiple spells in an analysis of poverty persistence, with half of those who end poverty spells returning to poverty within four years. Accounting for multiple spells shows that approximately 50 percent of blacks and 30 percent of whites falling into poverty in some year will have family income below the poverty line in five or more of the next ten years.

I. Introduction

Questions regarding the persistence of poverty are central to much past and current debate on the extent of poverty and public policies to address it. Early discussions of a "culture of poverty" implicitly assumed that a sizable portion of the poverty population remained in poverty for many years. More recent discussions of the underclass and of long-term welfare dependency also presume that some individuals remain poor over much of their lifetimes. Issues of income dynamics have been frequently studied by economists, but there are additional reasons for focusing on dynamics at the bottom of the income distribution. If poverty persists for many years, policymakers and others have good reasons for concern over the

Ann Huff Stevens is a professor of economics at Yale University. *This paper has benefitted from many helpful discussions with and comments from Gary Solon, Charles Brown, David Lam, Robert Moffitt, Rebecca Blank, and Marianne Page. Valuable feedback was provided by seminar participants at the University of Michigan, Yale University, and the National Bureau of Economic Research. The data used in this article can be obtained from the author beginning January 2000 through December 2003, at the Department of Economics, Box 208269, Yale University, New Haven, CT 06520.*
[Submitted February 1997; accepted June 1998]

consequences of such long-term deprivation. In addition, since government programs frequently provide assistance to those in poverty, it is important to document the extent to which certain individuals remain in poverty, and so remain eligible for public assistance, year after year. While interest in income and poverty dynamics has a long history, one very basic question has not been answered: how long will an individual falling into poverty spend below the poverty line? This question is central to our understanding of the concentration of poverty and of the degree of mobility in the lower portion of the income distribution.

This paper contributes to the literature on poverty dynamics in two ways. First, it provides estimates of the time spent in poverty over multiple spells. Previous spell-based measures of poverty persistence have been limited to single spells of poverty, and so have significantly understated total time spent below the poverty line. I estimate poverty persistence using two sets of hazard rates for movements into and out of poverty, controlling for observable characteristics and for unobserved heterogeneity in exit rates. This technique produces estimated distributions of time spent in poverty for individuals just beginning a poverty spell, with a variety of individual and household characteristics. Second, I supplement these estimates with results based on another approach common in the income dynamics literature involving the estimation of components-of-variance models. I then compare the results from these two models to each other and to those from a simpler method of directly tabulating years poor from panel data. These different models of poverty dynamics have rarely been estimated together, and the two approaches used here have not been compared in terms of their predictions of time spent in poverty.

II. Previous Research on Poverty Persistence

Previous authors have used several different approaches to study the dynamics and persistence of poverty.[1] Initial research into this area made use of newly available longitudinal data to observe and count the number of years individuals spend poor out of a fixed sample period.[2] The primary shortcoming of this strategy is that people who end (or begin) the period in poverty may be starting (or ending) a long stay in poverty, despite the fact that they appear to be poor in only one or two of the observed years. This leads to an understatement of persistent poverty, since some of those observed to be poor only a short time are actually in the midst of lengthy poverty spells that are censored by the beginning or end of the sample frame. In addition, these methods do not lend themselves to multivariate analyses of factors affecting the persistence of poverty.

Another commonly used method in the literature on income and poverty dynamics involves the estimation of components-of-variance models to describe the evolution of earnings or income over time. Lillard and Willis (1978) first applied this method

1. These methods have previously been summarized by Bane and Ellwood (1986) and Duncan and Rodgers (1991) among others.

2. Duncan et al. (1984), for example, used PSID data on individuals from 1969 through 1978 to count the number of years in which household income fell below the poverty line. They found that 33 percent of the sample experienced poverty in at least one year, but only about 5 percent were ''persistently'' poor, defined as being in poverty for eight out of the ten years.

to study poverty in their examination of earnings mobility in a sample of male household heads. Using their estimates of the permanent and transitory variance components of male earnings, Lillard and Willis derived probabilities of various time sequences of poverty or low-earnings status. Because their study was limited to the earnings of male heads of households, the authors noted that a more complete analysis of poverty within their framework would require incorporating family income (rather than male earnings) and changes in family composition over time, along with other factors.

Since the Lillard and Willis study, many authors have used variations of the permanent-transitory approach to study questions relating to poverty.[3] Only rarely, however, have these methods been applied to measures of family income or the family income-to-needs ratio used in conventional definitions of poverty in the United States.[4] In addition, many applications have continued to focus on male heads of household, or two-parent households, and thus missed much of the long-term poor population. Finally, though the parameters of these models can be used to derive probabilities of spending any number of years in poverty, this has not been done in most applications. The estimated variance components are often used to distinguish "permanent" from "transitory" poverty, rather than estimating distributions of time spent below the poverty line.

A different approach to the measurement of poverty persistence was used by Bane and Ellwood (1986). These authors used a spell-based, or hazard rate approach in their work. They focused on individual spells of poverty—consecutive years in which total household money income was less than the poverty line—and estimated the probability of ending these poverty spells, allowing for duration dependence in the hazard rate. This study offered several advantages over previous research. First, right-censored spells, those in progress at the end of the sample period, are easily incorporated into hazard rate models. Persons who remain in poverty through the end of the sample contribute to the estimation of exit probabilities (through the denominator of the exit rate) in all years through the censored year. Second, the spells approach easily incorporates poverty transitions involving changes in household structure, and provides a way to highlight the events associated with transitions into and out of poverty. Third, while Bane and Ellwood did not explore temporal changes in poverty exit rates, their approach is well suited to examination of the effects of time-varying factors on poverty persistence. Finally, Bane and Ellwood derived a variety of distributions based on the estimated exit rates, and clarified differences between the length of time in poverty for a person currently poor, and for someone just entering poverty. The findings of this study emphasized that, while most of those who fall into poverty will experience a short spell, the bulk of those currently poor are in the midst of a lengthy stay in poverty.

Despite these advantages, a major drawback of the Bane and Ellwood study is its

3. Gottschalk (1982), for example, estimates the percentage of married males with permanently low earnings using similar techniques.

4. One example is Duncan and Rodgers (1991) who apply components-of-variance techniques to the income-to-needs ratios for samples of black and white children. Similarly Duncan (1983) estimates permanent-transitory models for six-year samples of men, women, and children. Neither of these studies uses the estimated parameters to derive distributions of time spent in poverty.

focus on single spells.[5] Particularly in the years just after an exit from poverty, individuals are likely to fall back below the poverty line. I find that half of all individuals ending a poverty spell in a given year will again have incomes below the poverty line within four years. This results in a very different distribution of total years spent below the poverty line than is implied by an analysis of single spells, as I show below.[6]

In this paper I begin by developing a discrete-time hazard framework to account for multiple spells of poverty, and use this framework to derive estimates of total time poor. Although either the spells framework or a components-of-variance model could serve as a starting point for this examination of poverty dynamics, I begin with the spells approach for several reasons. For consistency with ''official'' definitions of poverty, and to account for important differences in welfare across households of different sizes with the same total income, I use the income-to-needs ratio, or total household income relative to a size-adjusted minimum needs standard. While components-of-variance models have been widely used and tested for earned income,[7] it is less clear which, if any, existing models are applicable to the income-to-needs ratio. Human capital theory provides some guidance in developing models for earned income; these theories are less directly applicable to the dynamics of household income-to-needs. Additionally, the spells approach easily allows for estimation over a variety of age groups and household structures, and for changes in household structure over time. Variance components models have typically been estimated on homogeneous groups, and on households with unchanging composition. For these reasons I begin by extending the hazard rate approach to consider poverty persistence over multiple spells. Given the frequency with which variance components models appear in the earnings dynamics literature, however, I supplement the hazard rate analysis with predictions of poverty persistence based on a variance components model of household income-to-needs.

III. Estimation Strategy

A. Basic Hazard Rate Model

The strategy in this paper centers around the estimation of two hazard rates: one for leaving poverty, and one for returning to poverty. I restrict the scope of this paper to questions of the persistence of poverty for those who ever become poor, rather

5. Bane and Ellwood did adjust their calculations to take into account some of the multiple spells that result from very brief periods above or below the poverty line. This still left many individuals in their sample experiencing more than one spell in the observed time frame, however. They reported that 31 percent of individuals in their sample experienced a second spell of poverty. This underestimates the true number of repeated spells because it does not account for right-censoring at the end of the sample; individuals exiting poverty spells near the end of the sample period will not be observed reentering poverty before the end of the sample.

6. The importance of multiple spells has been noted in the literature on spells of welfare receipt. See, for example, Gottschalk and Moffitt (1994), or Ellwood (1986). The same point has not previously been made with respect to poverty spells.

7. Even with respect to earned income among adult males, there is controversy over which models best capture the dynamics. See Baker (1997).

than examining the incidence of poverty for the entire population.[8] For now, assume that I start with a sample of spells with known durations. Let the probability of exiting poverty in any given year be represented by a standard logit specification. The probability of ending a spell of poverty after d years is given by the hazard function, λ^p_{idt}.

$$(1)\quad \lambda^p_{idt} = \frac{\exp(y_{idt})}{1 + \exp(y_{idt})}$$

$$(2)\quad y_{idt} = \alpha^p_d + \beta^p X_{it}$$

The subscript i indexes individuals; t refers to calendar years, and d indexes duration of the current spell. The hazard given in (1) depends on characteristics summarized by y_{idt}, which consists of duration effects (α^p_d) and the effects of other variables in X_{it} that may vary across people and across time, including calendar year dummies. The superscript p distinguishes the hazard rate for ending a poverty spell from the hazard for ending a nonpoverty spell, shown below. The probability of observing a completed poverty spell of length d is

$$(3)\quad f(d) = \left[\prod_{s=1}^{d-1} \frac{1}{1 + \exp(y_{ist})}\right]\left[\frac{\exp(y_{idt})}{1 + \exp(y_{idt})}\right]$$

The first term in (3) represents the probability of not exiting in each of the years prior to the dth year, and the second term is the probability of ending the spell in year d. Some of the spells in the data will continue beyond the end of the sample period, and these are easily incorporated into this framework. The probability of observing a right-censored spell is

$$(4)\quad 1 - F(d) = \prod_{s=1}^{d} \frac{1}{1 + \exp(y_{ist})}$$

Once an individual has ended a poverty spell, the probability of returning to poverty over the next several years can be modeled in the same way. The hazard function for returning to poverty is similarly defined as λ^n_{idt}.

$$(5)\quad \lambda^n_{idt} = \frac{\exp(z_{idt})}{1 + \exp(z_{idt})}$$

$$(6)\quad z_{idt} = \alpha^n_d + \beta^n X_{it}$$

The d subscript now refers to the duration of the spell out of poverty. The hazard for returning to poverty depends on z_{idt}, which again consists of the effects of spell

8. In theory, I could also estimate rates of first entry (as opposed to reentry) into poverty, but have chosen not to include first entry rates in this study. Including these rates would require that I observe individuals from birth forward to identify their initial entry into poverty. This means that I could, at most, directly estimate first-entry rates only up to age 22 (the length of the panel). Further, these estimates would be based on very small samples of individuals born into the survey since 1967.

duration (α_d^n) and of other individual and household characteristics (X_{it}). The probability of a spell out of poverty lasting for d years is

$$(7)\quad g(d) = \left[\prod_{s=1}^{d-1} \frac{1}{1 + \exp(z_{ist})}\right]\left[\frac{\exp(z_{idt})}{1 + \exp(z_{idt})}\right]$$

The contribution of a right-censored spell to the likelihood function is

$$(8)\quad 1 - G(d) = \prod_{s=1}^{d} \frac{1}{1 + \exp(z_{ist})}$$

These specifications are initially used to estimate, separately, exit rates from poverty and reentry rates into poverty.

The above discussion assumes that the duration of all spells is known. This will not be the case for left-censored of spells of poverty (and nonpoverty), or those already in progress at the start of the sample period. If the true model were driven only by spell duration, with no role for unobserved heterogeneity or omitted variables, estimation using only spells beginning after the start date of the sample would give consistent estimates of poverty transition rates for the population (Heckman and Singer 1984). In the presence of unobserved heterogeneity, however, eliminating spells in progress at the start of the sample, will induce a form of selection bias. Individuals who begin new spells after 1967 are likely to have higher transition probabilities (because they have experienced at least one transition since 1967) than the overall population. Some of those poor in 1967 will eventually be included in the analysis; if an individual who is poor in 1967 exits an initial spell, he or she will then be included in the estimation of the reentry hazard. Those eventually beginning a new spell of poverty will be included in the exit rate estimation. Individuals who are poor in 1967 and remain poor for many years, however, are selected out of the sample entirely. This may result in over-estimation of exit rates from poverty at long durations if the left-censored spells are not incorporated.

In results presented below I impose restrictions on the duration specification that allow me to include left-censored spells using an analysis sample beginning in 1973. These results show that the inclusion of left-censored spells makes little difference to the estimated exit and entry rates. More formal techniques also exist for dealing with left-censored spells. Moffitt and Rendall (1995), for example, include left-censored spells of female headship in their analysis by integrating over all possible (unobserved) histories up to the point of censoring. Several considerations lead me not to pursue this strategy here. First, it is unlikely that a model including both the left-censored spells (and integrating over all possible histories) and unobserved heterogeneity (discussed below) would be tractable. A related issue is that the model estimated by Moffitt and Rendall conditions the hazard rates only on age and calendar year (along with duration), both of which can be completely "backcast" and so contribute the necessary presample information to the likelihood function. Most of the models I estimate include, in addition to unobserved heterogeneity, variables such as sex and education of the household head that change over time and that would be more difficult to specify for the left-censored spells.

B. Unobserved Heterogeneity

The model above assumes that, controlling for duration and various observable characteristics, individuals' exit probabilities are independent across years. This will not be true, however, in the presence of unobservable person-specific characteristics that affect mobility out of and back into poverty. Further, once I allow for multiple spells, it may be particularly important to control for unobservable characteristics that influence individuals' transition rates. Allowing for correlation across an individual's exit and reentry probabilities over time may be necessary to accurately estimate total years spent in poverty over an extended time period. A person who has previously experienced a long poverty spell and then reenters poverty, for example, may be more likely to experience a long second spell. It may also be the case that reentry rates are related to exit rates from poverty; persons with particularly high exit rates from poverty may also have lower reentry rates.

To better capture the correlation across individual spells, I use a method suggested by Heckman and Singer (1984) and specify a likelihood function that allows for correlation across an individual's spells in and out of poverty. Although this approach has appeared frequently in the literature, (see, for example, Ham and Rea 1987) one additional complication here is that the heterogeneity components must enter two types of spells.[9] I allow heterogeneity to enter the hazard function through intercept terms in y and z above. Equations 2 and 6 above are modified to include an individual specific effect:

$$\text{(2a)} \quad y_{idt} = \theta_i^p + \alpha_d^p + \beta^p X_{it}$$

$$\text{(6a)} \quad z_{idt} = \theta_i^n + \alpha_d^n + \beta^n X_{it}$$

I assume that θ^p and θ^n can be characterized by discrete distributions with two support points each. There are two types of people with respect to spells of poverty. Individuals of Type 1 with respect to poverty spells occur in the population with some probability and have an intercept term given by θ_1^p; similarly, individuals of Type 2 have an intercept term θ_2^p. I also allow for two types of people with respect to spells out of poverty. Some proportion of the population have an intercept term in Equation 6a above of θ_1^n; the remaining proportion of the population have an intercept of θ_2^n.

Given this structure for the heterogeneity across spells in and out of poverty, the population of individuals ever becoming poor is characterized by the joint distribution of θ^p and θ^n. The joint distribution describing this heterogeneity is summarized in the two-by-two table below. R_1 through R_4 are the probabilities of observing each of the four types of individuals, with $R_1 + R_2 + R_3 + R_4 = 1$ where

	θ_1^p	θ_2^p
θ_1^n	R_1	R_2
θ_2^n	R_3	R_4

9. Eberwein, Ham, and LaLonde (1997) use a similar model for spells in and out of employment.

Each person is characterized by the vector

$$(9)\quad \underline{\theta}_i = \begin{pmatrix} \theta_i^p \\ \theta_i^n \end{pmatrix}$$

Given the joint distribution specified above, the vector $\underline{\theta}_i$ takes on four possible values in the population. Now, the contribution to the likelihood function conditional on being one of these four types, $\underline{\theta}_k$, can be written as

$$(10)\quad L_i(\underline{\theta}_k) = \left[\prod_{s=1}^{n1i} f_{is}(d;\underline{\theta}_k)\right]\left[\left(\prod_{s=1}^{m1i} g_{is}(d;\underline{\theta}_k)\right)\right][1 - F_i(d;\underline{\theta}_k)]^{\phi}[1 - G_i(d;\underline{\theta}_k)]^{1-\phi}$$

The term $n1i$ represents the number of non-right-censored spells of poverty an individual contributes to the function; $m1i$ is the number of non-right-censored spells out of poverty. ϕ is equal to one if the individual's history ends with a spell of poverty in progress and equals zero for individuals who end the period in a right-censored spell out of poverty. Summing over these conditional probabilities, each individual contributes to the likelihood function

$$(11)\quad L_i = \sum_{k=1}^{4} R_k L_i(\underline{\theta}_k)$$

The likelihood function described by (10) and (11) can now be maximized with respect to α_d, β, R_k and $\underline{\theta}_k$.

C. Interpretation of Transition Probabilities

These methods will produce estimates of transition rates out of and into poverty, but it is also important to interpret these estimates in terms of their implications for total time spent below the poverty line. For this I turn to microsimulation techniques, similar to those used in Moffitt and Rendall (1995) and Gottschalk and Moffitt (1994). The logit hazard rates specified above can be expressed in terms of the underlying latent index functions.

$$(12)\quad I^p = \theta_i^p + \alpha_d^p + \beta^p X_{it} + \varepsilon^p$$

$$(13)\quad I^n = \theta_i^n + \alpha_d^n + \beta^n X_{it} + \varepsilon^n$$

The terms I^p and I^n represent latent variables such that an exit from the indicated state (poverty or nonpoverty) occurs if the value is greater than zero, and no exit occurs otherwise. The error terms, ε^p and ε^n are assumed to be distributed independently and to follow the logistic distribution. Each individual has a fixed value of $\underline{\theta}_k$, distributed across the population according to the estimated probabilities R_k.

Given estimates of θ, R, α, and β, the above equations can be used to simulate the distribution of time spent below the poverty line for a group of individuals just falling into poverty. The error terms are generated by random draws from the logistic distribution. Using the estimated coefficients on observable characteristics, multiyear sequences of poverty status are simulated for 10,000 individuals just beginning a poverty spell. An exit occurs if I^p is greater than zero, and in the next period the possibility of returning to poverty is simulated in the same manner.

IV. Data and Samples

The data for this study come from the 1968 through 1989 interview years of the Panel Study of Income Dynamics, and correspond to calendar years 1967 through 1988.[10] The PSID collects data on household and individual earnings, income, family composition and labor market status, along with other characteristics. All persons are included in the analysis in each year they are present in the sample, including those who leave prior to the end of the survey.[11] Sample sizes are shown in Table A1. Because I use only years 1973 through 1988 in the analysis including left-censored spells, sample sizes for these years are shown as well. The unit of analysis in the hazard rate estimation is the individual, although the poverty definition relies on each individual's household income-to-needs level. This is necessary to allow for individuals to move from one household situation to another over time. Table A2 provides sample means for the individuals in spells of poverty and spells out of poverty.[12]

The definition of poverty follows the official Census Bureau poverty line. Individuals are considered poor in any year that their total household money income is less than the appropriate needs standard for their household size. The needs standard for the poverty cutoff is defined as 1.25 times that used in the Census Bureau poverty guidelines. This definition is commonly used by researchers working with PSID data and is meant to account for the consistently higher reported income in the PSID than in Census Bureau data.

Finally, as is well known, the PSID is not a random sample; the original survey oversampled low-income households. This selection based on poverty status means that all estimates must be weighted to consistently estimate population parameters. All of the hazard model results reported below are weighted using the PSID sample weights. This weighting, and the clustered sample design of the PSID, mean that traditional standard error estimates will be biased. To correct for this bias, standard errors are estimated using the method of balanced repeated replications (see Kish and Frankel 1970); these methods are discussed in more detail in the Appendix. Another source of bias in traditional standard error estimates is that I use multiple

10. For most of the results presented, only data for calendar years 1969 forward are included in the estimation. Data for the first two years of the sample are not used since, initially, only spells starting after the beginning of the sample are included. Left-censored spells are later incorporated, using only the data from 1973 forward.

11. One group not included in the sample is individuals joining the survey after 1967 through marriage or other movements into existing survey households. These individuals are considered "non-sample" persons by the PSID. Since the PSID sample and weights are designed to be representative without including these nonsample persons, I do not include them in the analysis. In contrast, people who are born to original PSID sample members are included.

12. To interpret the means in Table A2 correctly, note that they are calculated over all person-years in and out of poverty. Average income in poverty spells, for example, represents the average over the stock of all individuals currently poor. Individuals in spells of poverty contribute to this average in each of the years they remain poor.

individuals from the same household and so observations within households are not independent. The techniques to address problems associated with the clustering induced by the survey design will also correct for this lack of independence within families.

V. Results from the Hazard Rate Model

A. *Without Unobserved Heterogeneity*

To establish the broad patterns of exit from and reentry into poverty, I first present simple estimates of the exit and reentry rates controlling only for duration of the spell and race.[13] Table 1 shows hazard rates based on the logit models including only duration terms, estimated separately for blacks and whites.[14] As expected, both exit and reentry rates decline substantially with duration. The probability of ending a poverty spell after one year poor is .53; after four years in poverty the exit rate is .23. The reentry rates show the substantial risk of returning to poverty and the importance of multiple spells. Among blacks, the probability of returning to poverty after one year out is more than a third. Even through the sixth year out of poverty, blacks have reentry rates in excess of .10. Differences by race show the greater persistence of poverty among blacks, as the consequence of both lower exit rates and higher reentry rates. Standard error estimates in Table 1 (and in Tables 2 through 5) are calculated using balanced half-sample techniques, and are more fully described in the Appendix.[15]

A more useful way to view the transition rate estimates is in terms of their implications for the number of years spent poor. The estimated transition rates from Table 1 are used to simulate the number of years spent poor over the next ten years. The first column of Table 2 shows the distribution of years spent poor in single spells of poverty, calculated using only the exit rates, and not taking multiple spells into account. Column two uses both the exit and reentry rates to simulate years spent poor over multiple spells. These table entries are probabilities that a person just falling into poverty will be poor for the given number of years out of the next ten. The final column of Table 2 comes from observing cohorts of persons in the data who enter poverty during each of the years from 1968 through 1979, and counting years poor over the next ten years. This is a slight modification of the technique of directly tabulating observed years in poverty discussed in Section II, conditioning

13. The results presented here are for "whites" and "blacks." The black subsample includes very small numbers of other races as well. The results are not sensitive to dropping these individuals from the sample, or, alternatively, including them in the whites category. Sample sizes are too small for a separate analysis of this group.

14. Estimation with only the duration terms produces estimates identical to those resulting from nonparametric estimation in which the number of persons ending a spell after exactly d years poor is divided by the total number poor (or nonpoor for the reentry hazard) for at least d years.

15. The standard errors resulting from this method are typically two to two and a half times as large as those calculated without correcting for the nonindependence of observations. This ratio is larger than that typically reported by individuals correcting for "design effects" using the PSID data (see Hill 1981, for example) because use of multiple family members from the same household causes an additional source of "clustering" and an additional source of bias in conventional standard error estimates.

Table 1
Exit and Reentry Rates by Duration and Race (1969–88)

Duration (years)	Exit Rates All	Exit Rates Blacks	Exit Rates Whites	Reentry Rates All	Reentry Rates Blacks	Reentry Rates Whites
1	0.53	0.42	0.57	0.27	0.35	0.23
	(0.011)	(0.017)	(0.013)	(0.010)	(0.017)	(0.011)
2	0.36	0.32	0.38	0.16	0.23	0.14
	(0.014)	(0.021)	(0.019)	(0.010)	(0.020)	(0.011)
3	0.27	0.20	0.31	0.11	0.13	0.10
	(0.016)	(0.017)	(0.025)	(0.009)	(0.013)	(0.012)
4	0.23	0.17	0.27	0.09	0.11	0.08
	(0.014)	(0.026)	(0.022)	(0.007)	(0.016)	(0.010)
5	0.19	0.15	0.22	0.08	0.10	0.07
	(0.020)	(0.022)	(0.028)	(0.008)	(0.015)	(0.009)
6	0.16	0.13	0.19	0.07	0.12	0.05
	(0.018)	(0.024)	(0.024)	(0.008)	(0.025)	(0.011)
7	0.15	0.15	0.15	0.06	0.05	0.06
	(0.029)	(0.036)	(0.040)	(0.009)	(0.011)	(0.011)
8	0.13	0.08	0.17	0.04	0.04	0.04
	(0.022)	(0.014)	(0.045)	(0.010)	(0.008)	(0.013)
9	0.12	0.08	0.16	0.05	0.07	0.05
	(0.022)	(0.022)	(0.046)	(0.009)	(0.029)	(0.009)
10	0.11	0.10	0.12	0.04	0.04	0.04
	(0.015)	(0.018)	(0.035)	(0.004)	(0.007)	(0.005)

Note: Standard error of hazard rate in parentheses.

on entry into poverty to avoid problems with censoring of spells in progress. These direct tabulations provide a useful benchmark for evaluating the accuracy of the hazard model.

The distributions in columns one and two of Table 2 illustrate the difference between single and multiple spell measures of the persistence of poverty. The distribution of single spell lengths suggests that a considerably larger proportion of the population will experience short stays in poverty than do the other columns. Comparing columns two and three shows the extent to which the simple hazard models capture the observed distribution of years poor following entry. The simulated distributions of years poor over multiple spells differ from those based on directly tabulating years poor. In particular, the simulated distributions consistently underestimate the proportion of individuals who will be poor for very few years. This is not surprising, given the absence of controls for any observables beyond spell duration and the simple manner in which multiple spells have been combined. By controlling for important observable characteristics, as well as for unobserved heterogeneity, in the following sections, the hazard model can better capture observed patterns. Despite

Table 2
Distribution of Years in Poverty—Single Spells and Multiple Spells

Years in Poverty	Distribution of Single Spell Lengths	Percent of Years Poor Out of Next Ten—Simulated	Percent of Years Poor Out of Next Ten—Actual
		Total Sample	
1	52.5	17.5	29.5
2	17.2	15.5	14.7
3	8.2	13.9	11.7
4	5.0	11.9	9.3
5	3.2	9.8	6.2
6	2.2	7.8	7.3
7	1.8	6.4	5.5
8	1.2	5.2	5.0
9	1.0	4.4	5.7
10	7.7	7.6	5.2
	100.0	100.0	100.0
		Blacks	
1	41.8	9.9	17.7
2	18.5	9.8	10.8
3	8.0	10.1	10.5
4	5.4	10.6	7.5
5	3.9	9.9	6.8
6	2.9	9.5	9.4
7	2.9	9.3	7.4
8	1.4	8.8	8.1
9	1.2	7.8	10.7
10	14.0	14.0	11.0
	100.0	100.0	100.0
		Whites	
1	57.0	21.3	33.8
2	16.6	18.1	16.0
3	8.2	15.2	12.2
4	4.9	12.1	10.0
5	3.0	9.4	5.9
6	1.9	6.9	6.5
7	1.3	5.1	4.8
8	1.2	3.9	3.9
9	0.9	3.2	3.9
10	5.0	5.0	3.9
	100.0	100.0	100.0

this divergence, however, the picture of total time spent poor based on the multiple spell exit and reentry rates is a much better predictor of observed poverty persistence than are the results based on single spells. The means of the distributions in Columns 2 and 3 are 4.4 and 4.0 years, compared to a mean duration of single spells of just 2.7 years.

To introduce additional covariates into the hazard functions, I next estimate the exit and re-entry rates using the logit framework described above. Tables 3 and 4 contain the estimated coefficients from the logit estimation. These estimates are based on models that include controls for calendar year, age, education of the household head, and sex of the household head, in addition to the duration terms.[16] In PSID terminology the male is the default ''head of household'' and so a female head in this context is equivalent to a *single* female head of household. I initially allow separate duration effects for durations of one to nine years, and another for durations of ten years or more. The predicted hazard rates based on these logit estimates can be derived by substituting the estimated coefficients into the hazard functions given in Equations 1 and 5.

The results in Tables 3 and 4 show that characteristics of individuals and their families affect the probabilities of exiting and returning to poverty in predictable ways. The education level of the household head and whether the household is headed by a single female have large impacts on the exit and reentry rates. The effect on exit rates of being in a female-headed household is summarized by a logit coefficient of $-.77$ for the overall sample. This translates into hazard rates for leaving poverty after one year of .39 for those in female-headed households, compared with .58 for those in households with male heads. The effect of female headship on reentry probabilities is large as well. After one year out of poverty, those in households headed by women have a reentry probability of one-third; the comparable figure for those in male-headed households is just under one-fifth. The impact of female headship and educational levels are more fully illustrated below using microsimulations to derive distributions of time in poverty by individuals with different characteristics.

The coefficient estimates also illustrate age-related patterns of poverty transitions. Exit rates from poverty are lowest for young children and peak at ages 18 through 24, reflecting the fact that many poverty transitions involve household structure changes that are common during early adulthood. Exit rates fall for individuals aged 55 and over. Reentry rates follow the reverse pattern, reaching a minimum for adults ages 25 to 34.

Variation in the hazard rates by year is also shown in Tables 3 and 4. The year-to-year differences in the hazard rates partially reflect changes in overall economic conditions. During the recession of 1975, for example, the exit rate falls from .62 to .51 and the reentry rate is higher than average. Such variations across years are due to both business cycle influences and longer term trends. Over time, there has been a slight downward trend in exit rates, controlling for aggregate business cycle conditions. For example, the rate of GNP growth was 2.5 percent in 1979 and a slightly higher 2.7 percent in 1986; the exit rates from poverty (at a duration of 1

16. Models were also estimated with a variety of interaction terms, as well as a few additional characteristics; estimated coefficients for these terms were generally not significant.

Table 3
Exit Rate Logit Coefficients by Race

	All	Blacks	Whites
Duration			
1	0.32 (0.14)	0.08 (0.23)	0.40 (0.15)
2	−0.32 (0.18)	−0.28 (0.28)	−0.35 (0.19)
3	−0.76 (0.16)	−0.83 (0.26)	−0.73 (0.17)
4	−0.94 (0.19)	−1.06 (0.33)	−0.87 (0.19)
5	−1.13 (0.18)	−1.21 (0.28)	−1.07 (0.21)
6	−1.29 (0.20)	−1.34 (0.29)	−1.19 (0.22)
7	−1.34 (0.23)	−1.20 (0.31)	−1.44 (0.31)
8	−1.49 (0.24)	−1.87 (0.30)	−1.21 (0.36)
9	−1.56 (0.28)	−1.82 (0.37)	−1.31 (0.44)
10 or more	−1.68 (0.21)	−1.63 (0.30)	−1.64 (0.39)
Education > = 12	0.39 (0.07)	0.44 (0.11)	0.33 (0.08)
Female head	−0.77 (0.07)	−0.81 (0.11)	−0.63 (0.08)
Age			
0–5	−0.43 (0.06)	−0.48 (0.11)	−0.33 (0.06)
6–12	−0.20 (0.06)	−0.24 (0.06)	−0.19 (0.07)
13–17	−0.15 (0.08)	−0.12 (0.08)	−0.15 (0.11)
18–24	0.22 (0.05)	0.28 (0.10)	0.20 (0.24)
34–44	0.03 (0.09)	0.16 (0.14)	−0.03 (0.09)
45–54	0.12 (0.07)	0.07 (0.12)	0.11 (0.10)
55 or older	−0.19 (0.08)	0.03 (0.16)	−0.35 (0.10)
Year			
69	−0.21 (0.40)	0.72 (0.38)	−0.72 (0.57)
70	0.24 (0.25)	0.49 (0.39)	0.13 (0.31)
71	0.49 (0.23)	0.48 (0.36)	0.49 (0.28)
72	0.31 (0.20)	−0.20 (0.36)	0.47 (0.21)
73	0.55 (0.20)	0.32 (0.35)	0.69 (0.24)
74	0.23 (0.20)	0.19 (0.29)	0.26 (0.25)
75	−0.23 (0.18)	−0.17 (0.30)	−0.24 (0.20)
76	0.37 (0.21)	0.11 (0.36)	0.52 (0.23)
77	0.02 (0.21)	−0.23 (0.27)	0.14 (0.21)
78	0.20 (0.18)	0.17 (0.29)	0.23 (0.20)
79	0.27 (0.20)	0.23 (0.25)	0.29 (0.26)
80	−0.34 (0.20)	−0.23 (0.31)	−0.40 (0.22)
81	−0.24 (0.19)	−0.73 (0.32)	−0.06 (0.23)
82	−0.39 (0.16)	−0.47 (0.25)	−0.36 (0.17)
83	−0.22 (0.15)	−0.48 (0.24)	−0.11 (0.16)
84	0.12 (0.19)	−0.02 (0.26)	0.19 (0.20)
85	−0.21 (0.15)	−0.33 (0.23)	−0.14 (0.21)
86	−0.00 (0.19)	−0.31 (0.29)	0.13 (0.19)
87	−0.10 (0.18)	−0.36 (0.31)	0.01 (0.18)

Note: Standard errors of logit coefficients in parentheses.

Table 4
Reentry Rate Logit Coefficients by Race

	All	Blacks	Whites
Duration			
1	−1.43 (0.09)	−1.16 (0.22)	−1.57 (0.13)
2	−2.07 (0.13)	−1.69 (0.23)	−2.25 (0.15)
3	−2.42 (0.13)	−2.29 (0.23)	−2.49 (0.17)
4	−2.66 (0.12)	−2.48 (0.23)	−2.74 (0.18)
5	−2.80 (0.10)	−2.50 (0.18)	−2.93 (0.15)
6	−2.96 (0.13)	−2.36 (0.23)	−3.25 (0.23)
7	−3.05 (0.17)	−3.23 (0.24)	−3.03 (0.20)
8	−3.38 (0.20)	−3.34 (0.22)	−3.41 (0.26)
9	−3.13 (0.16)	−2.77 (0.44)	−3.30 (0.19)
10 or more	−3.49 (0.13)	−3.44 (0.20)	−3.53 (0.16)
Education > = 12	−0.38 (0.06)	−0.56 (0.11)	−0.28 (0.08)
Female head	0.74 (0.07)	0.86 (0.09)	0.61 (0.08)
Age			
0–5	0.61 (0.08)	0.59 (0.10)	0.59 (0.11)
6–12	0.48 (0.07)	0.38 (0.09)	0.52 (0.09)
13–17	0.47 (0.08)	0.33 (0.15)	0.52 (0.09)
18–24	0.18 (0.06)	0.09 (0.10)	0.20 (0.09)
34–44	0.17 (0.07)	0.01 (0.15)	0.23 (0.08)
45–54	0.14 (0.10)	−0.16 (0.15)	0.26 (0.12)
55 or older	0.32 (0.07)	0.12 (0.11)	0.46 (0.08)
Year			
69	−0.97 (0.28)	−1.06 (0.35)	−0.89 (0.40)
70	−0.33 (0.24)	−0.35 (0.30)	−0.31 (0.40)
71	0.02 (0.24)	−0.06 (0.34)	0.08 (0.28)
72	−0.26 (0.15)	0.14 (0.25)	−0.49 (0.18)
73	−0.39 (0.14)	−0.45 (0.26)	−0.31 (0.20)
74	0.35 (0.14)	0.40 (0.21)	0.34 (0.20)
75	0.05 (0.13)	0.24 (0.23)	−0.07 (0.19)
76	0.00 (0.14)	0.07 (0.24)	−0.01 (0.19)
77	−0.11 (0.16)	0.27 (0.25)	−0.28 (0.19)
78	−0.09 (0.15)	0.03 (0.21)	−0.12 (0.17)
79	−0.11 (0.15)	−0.20 (0.25)	−0.07 (0.20)
80	0.15 (0.13)	0.06 (0.23)	0.18 (0.17)
81	0.21 (0.15)	0.14 (0.26)	0.25 (0.18)
82	0.32 (0.14)	0.38 (0.24)	0.31 (0.18)
83	0.15 (0.13)	0.33 (0.15)	0.08 (0.18)
84	0.03 (0.11)	0.43 (0.18)	−0.09 (0.14)
85	0.20 (0.13)	0.04 (0.24)	0.06 (0.16)
86	−0.06 (0.14)	−0.06 (0.25)	−0.06 (0.17)
87	−0.13 (0.16)	0.14 (0.19)	−0.21 (0.20)

Note: Standard errors of logit coefficients in parentheses.

year) in these years, however, were .64 and .58.[17] The reentry rates over time show less of a trend, and also display a weaker relationship with the state of the economy.

B. Left-Censoring

The results presented so far have been based on spells of poverty that begin after the start of the sample. As noted above, ignoring left-censored spells can introduce a source of bias into the hazard rate estimates. I next examine the sensitivity of these estimates to left-censoring of spells in progress at the start of my sample. Specifically, I modify the duration specification so that left-censored spells can be incorporated. First, I reduce the number of separate duration terms in the hazard rates from ten to six. Given the estimates of the duration terms reported in Tables 3 and 4, this is a reasonable restriction. Beyond six years, the hazard functions estimated separately for blacks and whites begin to fluctuate and the duration terms no longer decline monotonically. A Wald test confirms that for both blacks and whites the hypotheses of equality of the duration terms beyond six years cannot be rejected. Using this specification, I can now estimate the model using data from 1973 through 1988. In 1973, I observe whether each person has been poor in the previous six years (since 1967), and so can use all poverty spells currently in progress in the estimation, including those that began prior to the start of the sample and now have durations of six years or more.

Inclusion of the left-censored spells makes remarkably little difference to the estimated transition probabilities. The predicted exit rates for durations of six years or longer are .15 and .16 with and without the left-censored spells, with standard errors of approximately .01. Estimated re-entry rates with and without the left-censored spells of six years or more are .03 and .02.[18] Overall, the bias from omitting left-censored spells appears to be extremely small.[19] For the remainder of the paper I include the left-censored spells of poverty in the analysis and specify the duration terms in this way.

C. Correlation Across Spells

I next examine the correlation of exit and reentry rates across spells by incorporating unobserved heterogeneity into the exit and reentry hazards. The likelihood function given in Equations 9–11 is maximized with respect to the parameters of the hazard

---

17. These year effects can be more formally decomposed by regressing the estimated year effects on a trend term and a business cycle indicator. These results are discussed in Stevens (1994, 1995).

18. The left-censoring issues are slightly different in the case of spells out of poverty. The nonpoor population in 1967 consists of two unobservable groups—those previously poor and those never poor. By using all individuals not poor in 1967 (left-censored spells out of poverty), I include a large number of individuals with no previous poverty spells. Despite this over-correction in the case of left-censored spells out of poverty, estimated rates of entry into poverty are only modestly affected and I do not include the left-censored nonpoverty spells in the rest of the analysis.

19. An alternative examination of the importance of left-censoring was also performed. I used all nonleft-censored observations to estimate the model with ten duration terms allowed, along with age and calendar year variables. I then artificially left-censored this sample at 1973 and eliminated those spells that were in progress in 1973. The comparison of samples left-censored in 1968 and in 1973 showed virtually no effect of the additional left-censoring.

functions and the parameters of the discrete distribution of the heterogeneity terms. Table 5 shows the estimated parameters. Because I have included a full set of duration dummy variables, one of the support points of the heterogeneity distribution is normalized to zero. This model was estimated with and without calendar year effects; results shown in Table 5 do not include calendar year terms. Inclusion of year effects does not alter the estimated heterogeneity or other parameters.

Initially, as described in Section III, I allowed the heterogeneity terms with respect to exit and reentry to be freely correlated with one another, so that there were four possible types in the population. In the case of whites, the estimation indicated that the probabilities corresponding to two of the four cells in the table were virtually zero. I thus restricted the heterogeneity terms to allow for only two types of individuals in the case of whites.[20] When the parameters are not interior to the parameter space, the asymptotic normality of the maximum likelihood estimator can no longer be established, and so standard error estimates from the maximum likelihood estimation will be invalid (Heckman and Walker 1990). Standard errors for whites shown in Table 5, like those in previous tables, are based on balanced half-sample replications that do not rely on the estimator's asymptotic normality.

The estimated distribution of the heterogeneity terms has most of the mass of the distribution (96 percent) at one support point. Exit rates corresponding to this support point for the reference group (adults in households headed by males with less than a high school education) are .64 in the first year of a poverty spell and decline to around .28 after six or more years in poverty. Reentry probabilities for this group range from .17 in the initial year out of poverty to .04 after six or more years out. The other support point occurs with a probability of just four percent, and corresponds to very low probabilities of escaping poverty, ranging from .25 to .07 for the reference group. The probability of returning to poverty among individuals of this ''type'' is very high, starting at .86 in the initial year following a poverty spell, and declining only to .55 after six or more years.

The estimated heterogeneity parameters for blacks in Table 5 show a different pattern. Forty-four percent of blacks are estimated to have relatively high exit rates and low reentry rates. For this group, exit rates range from .63 after one year in poverty to .28 after six or more years; reentry rates range from .24 to .06. These rates are very close to the estimated transition rates for the bulk of the white population noted above. At the other extreme, 14 percent are estimated to have low rates of exit and high rates of reentry. Relative to whites, the heterogeneity distribution for blacks suggests that a substantially larger proportion of the black population falling below the poverty line (14 versus 4 percent) will face extremely persistent poverty. Another difference in poverty dynamics for blacks is that a third category makes up a substantial part of the overall population. Forty-two percent of blacks are estimated to have low exit rates, but also low rates of reentry. This group experiences long spells of poverty, but does not face the extremely high reentry probabilities consistent with permanent poverty. The estimated standard errors show that the

20. Cutler (1995) takes a similar approach in estimating a joint distribution of heterogeneity terms. Specifically, he estimates the joint distribution of heterogeneity terms correlated across equations for hospital mortality, hospital readmission, and nonhospital mortality. Cutler begins by allowing two support points in each of his three equations, but finds that only three of the eight potential combinations of these support points across the equations occur with a probability greater than zero.

Table 5
Coefficients from Maximum Likelihood Estimation of Hazard Rates with Unobserved Heterogeneity

	Blacks		Whites	
	Spells Out of Poverty	Spells in Poverty	Spells Out of poverty	Spells in Poverty
Duration (years)				
1	−1.15	−0.33	1.80	−1.08
	(0.02)	(0.03)	(0.25)	(0.18)
2	−1.56	−0.64	1.17	−1.78
	(0.17)	(0.06)	(0.22)	(0.19)
3	−2.12	−1.17	0.97	−2.16
	(0.12)	(0.31)	(0.26)	(0.21)
4	−2.32	−1.36	0.77	−2.24
	(0.03)	(0.07)	(0.23)	(0.19)
5	−2.24	−1.49	0.56	−2.36
	(0.08)	(0.16)	(0.28)	(0.25)
6 or more	−2.75	−1.78	0.20	−2.62
	(0.05)	(0.06)	(0.23)	(0.22)
Age				
<6	0.72	−0.51	0.43	−0.36
	(0.03)	(0.02)	(0.13)	(0.10)
6–17	0.36	−0.27	0.36	−0.14
	(0.03)	(0.02)	(0.06)	(0.06)
18–24	−0.09	0.25	0.10	0.20
	(0.01)	(0.05)	(0.07)	(0.05)
>54	−0.27	−0.14	0.28	−0.40
	(0.14)	(0.01)	(0.08)	(0.08)
Female head	0.88	−0.85	0.68	−0.74
	(0.06)	(0.10)	(0.08)	(0.10)
High school or more	−1.03	0.27	−0.26	0.32
	(0.16)	(0.26)	(0.09)	(0.06)
θ_1^p		0.85		1.66
		(0.10)		(0.20)
θ_1^n	1.80		−3.36	
	(0.41)		(0.21)	
θ_2^p		0		0
θ_2^n	0		0	
$R1$ (θ_1^p, θ_1^n)	0[a]		0.96	
			(0.04)	
$R2$ (θ_1^p, θ_2^n)	0.44		0[a]	
	(0.38)			
$R3$ (θ_2^p, θ_1^n)	0.14		0[a]	
	(0.22)			
$R4$ (θ_2^p, θ_2^n)	0.42		0.04	
	(0.43)		(0.04)	

Note: Standard errors in parentheses.
a. Probabilities for four support points were initially included in the model. The indicated probabilities, however, were estimated to be zero. The estimation was then performed restricting the probabilities of these "types" to be equal to zero.

probabilities associated with the heterogeneity distributions are very imprecisely estimated.

For the results for blacks in Table 5 the estimation of standard errors using the balanced half-sampling procedure was not successful. To obtain an efficient estimate of these standard errors requires performing the maximum likelihood estimation 32 times on selected half-samples. In the case of blacks, convergence was not achieved in nine of the 32 half-sample estimates. As an alternative, the standard errors reported in Table 5 are based on half-sample parameter estimates across only eight randomly chosen half-samples. As shown in Wolter (1985), any particular half-sample estimate of a parameter, β_k, can be used to form an approximately unbiased estimate of the variance of β (the full sample parameter estimate) by taking the squared deviation of β_k from β. Such variance estimates will not be as precise as those based on the fully balanced replication procedure and the standard errors reported for blacks in Table 5 should be interpreted with caution.

Finally, I also attempted to allow additional support points in the heterogeneity distribution, proceeding in two steps. First, I separately estimated hazards for spells in poverty and for spells out of poverty, increasing the support points for the heterogeneity distributions in each type of spell from two to three. In the hazard for ending poverty spells, inclusion of three support points slightly improved the maximum of the likelihood function, for both whites and blacks. Second, I used the three support points from the separate poverty spell estimation (and two support points from the nonpoverty spells) as starting values in the joint estimation of poverty and nonpoverty spells. This estimation was unsuccessful, however. These models either failed to converge or resulted in estimated probabilities of zero for the additional combinations of support points.

Given the estimation difficulties associated with the non-parametric heterogeneity distributions, I have also examined the sensitivity of these results to the inclusion of unobserved heterogeneity and to alternative specifications of the relationships across individual spells. One alternative method to specify relationships across separate spells is to include terms for lagged duration dependence. Results from estimation of the exit and reentry hazards that control for the length of previous spells confirm that those experiencing long stays in poverty are more likely to return to poverty. Among blacks, for example, approximately half of all individuals ending a short poverty spell remain out of poverty for at least five years. For those experiencing longer spells, however, only 37 percent stay above the poverty line for at least five years. In addition to influencing reentry probabilities, the length of the poverty spell also affects the duration of subsequent spells for those experiencing two or more spells of poverty. The expected poverty spell length for individuals with a previous spell of less than five years is 2.3 years for whites and 3.9 years for blacks. For those with a previous spell of five years or longer, the expected durations of poverty spells for whites and blacks are 4.8 and 5.6 years.

It is generally not possible to separately identify heterogeneity from structural duration dependence or lagged duration dependence. My motivation for estimating these models with both the heterogeneity and lagged duration dependence is to accurately capture patterns of poverty persistence. No structural interpretation for either the heterogeneity or the duration dependence is intended. To examine whether the choice of how to model correlation across spells affects the estimated persistence of poverty, I have simulated distributions of time in poverty using parameters from three different specifications: models based on individual spells of poverty with no unobserved heterogeneity (summarized in Tables 3 and 4); those including unobserved heterogeneity (Table 5); and those described above including terms for length of previous poverty spells and previous spells out of poverty. The three sets of parameters resulted in virtually identical distributions of time spent in poverty. Neither the unobserved heterogeneity distributions nor the presence of lagged duration dependence substantively changed my predictions of time spent in poverty, conditioning on (current) spell duration and a small set of observable characteristics. The estimated times in poverty presented in the remainder of the paper are for those based on the parameters shown in Table 5, including the unobserved heterogeneity parameters.

Table 6 illustrates how poverty persistence varies with age, race, and the household head's level of education. For two age groups beginning a stay in poverty at birth, and at age 20, mean time in poverty and the percentage spending more than half of the next ten years poor are shown by sex and education of the household head. Among children, shown in the upper portion of the table, for example, the mean years poor varies from 3.3 out of the next ten for whites in households headed by a male with at least a high school education to 8.3 of the next ten for blacks in households headed by a female with less than a high school education. Poverty is an extremely persistent state for those in female-headed households, and for those in households where the head does not have a high school education. Nearly 90 percent of black children in the low-education, female-head category who fall below the poverty line will be poor for six or more of the next ten years.

The effects of female headship shown in Table 6 are large, and this is partially a function of the simulation being performed in this table. The contrast between male- and female-headed households captures differences between individuals remaining in a male- or female-headed household over the entire ten year period. In the case of female headship, in particular, holding constant headship status over such a lengthy period may not be realistic.[21] Distributions can also be calculated based on alternative profiles of female headship over the period. Suppose, for example, that a child is born into poverty and spends only the first three years of life in a female-headed household, and the following seven years in a two parent household. The estimated average time spent in poverty in this case is 5.2 years for whites (where the head has less than a high school education) and seven years for blacks. This can be compared with four years (for whites) and six years (for blacks) among children who spend no time in a female-headed family. Even a short time period spent in a female-headed household significantly increases poverty persistence.

VI. A Components-of-Variance Approach to the Measurement of Poverty Persistence

As noted in Section II an alternative approach used to study poverty dynamics is the estimation of parameters of a time series process describing earnings

21. Persistence of female headship for ten years is unusual for white women, but not unusual among blacks. Among black female heads in my sample starting a poverty spell, more than half of those observed for the next ten years remain female heads the entire time.

Table 6
Expected Years in Poverty by Age and by Education and Sex of Household Head

	Blacks		Whites	
	Mean Years Poor Over Next Ten	Percent Poor More Than Five of Next Ten Years	Mean Years Poor Over Next Ten	Percent Poor More Than Five of Next Ten Years
Starting age = 1				
Male household head				
Less than high school	5.99	55.6%	4.04	27.6%
High school or more	4.35	33.3%	3.25	16.7%
Female household head				
Less than high school	8.30	89.5%	6.44	63.0%
High school or more	6.96	68.9%	5.41	47.3%
Starting age = 20				
Male household head				
Less than high school	3.93	26.4%	2.93	13.0%
High school or more	2.65	11.2%	2.42	7.7%
Female household head				
Less than high school	6.52	64.1%	4.92	39.6%
High school or more	4.81	39.1%	3.99	26.4%

or income dynamics. These estimated parameters can then be used to derive statements about poverty persistence. In this Section I implement such an approach and compare the results to those from the hazard model.

Components-of-variance models were first used to study poverty or low-earnings status by Lillard and Willis (1978). Since that time many variations of this basic approach have appeared in the literature. The implementation here differs from much previous work in that the dependent variable used is the log of the income-to-needs ratio, rather than individual earnings or income. Another difference from Lillard and Willis and some other authors involves the estimation method. Recently components-of-variance models have been estimated not by maximum likelihood, but by a generalized method-of-moments approach that minimizes the distance between the sample covariances and their population counterparts (Abowd and Card 1989, Baker 1997). The newer approach avoids the previous literature's assumption that the variance components are normally distributed. I use this approach and estimate the variance components using an equally weighted minimum distance estimator.[22]

Initially following Lillard and Willis' (1978) model for individual earnings, suppose that the evolution of the income-to-needs ratio over time can be summarized by the following equations:

$$(14)\quad Y_{it} = X_{it}\beta + \varepsilon_{it}$$
$$(15)\quad \varepsilon_{it} = \alpha_i + u_{it}$$
$$(16)\quad u_{it} = \rho u_{i,t-1} + \nu_{it}$$

Y_{it} is the log of the income-to-needs ratio; X_{it} contains terms for age and year effects (and potentially other variables); and the error term consists of a person-specific component (α_i) and a transitory term that follows an $AR(1)$ pattern.

The model summarized by (14)–(16) implies a particular form for the population covariances of ε_{it}. Specifically,

$$(17)\quad \text{cov}(\varepsilon_{it}, \varepsilon_{is}) = \text{var}(\alpha_i) + \frac{\rho^{|t-s|}}{1-\rho^2}\text{var}(v_{it}) = \sigma_\alpha^2 + \frac{\rho^{|t-s|}}{1-\rho^2}\sigma_v^2$$

The unique elements of the population covariance matrix can be represented by the column vector $f(\sigma_\alpha^2, \rho, \sigma_v^2)$. Let the elements of the estimated covariance matrix be similarly arrayed into a column vector c. This matrix is formed from the residuals of separate regressions by year of the log of the income-to-needs ratio on age and age squared, as well as race, education and sex of the household head (for comparability with the hazard model results).[23] The estimated covariance matrix is included in the appendix as Table A3. The parameters of the model are estimated by minimizing

$$(18)\quad (c - f(\sigma_\alpha^2, \rho, \sigma_v^2))'(c - f(\sigma_\alpha^2, \rho, \sigma_v^2))$$

with respect to σ_α^2, ρ and σ_v^2.

The estimated parameters of the error structure are shown in Table 7. The first model estimated is that specified above, with an individual effect and a transitory term that follows an $AR(1)$ process. One shortcoming of this model is that it predicts constant variances over time, a prediction that is inconsistent with the estimated covariance matrix. A simple way to accommodate the nonstationarity of the variances is to allow the transitory term (v_{it}) to have year-specific variances. The parameters of this model are shown in the second row of the table. The third and fourth

22. This estimator will provide consistent, but asymptotically inefficient, estimates of the parameters. The asymptotically efficient estimator uses the inverse of the consistently estimated variance-covariance matrix of the second moments as a weighting matrix. It has been common in the recent literature, however, to use the identity matrix (equal weights) instead. This is because of the potential singularity of the estimated variance-covariance matrix, and its poor small sample properties (Baker 1997). Abowd and Card (1989) show that the optimally weighted estimator may produce misleading estimates, particularly in the case of poorly fitting models.

23. The sample contains the nonrandom portion of the PSID and so both the first-stage regressions and the calculation of the covariances use sample weights. I have also performed the analysis using only the random sample portion of the PSID without weights and obtained similar results. Lillard and Willis (1978) and Baker (1997) use only the random portion of the PSID. Abowd and Card (1989) show results for samples with and without the non-random portion, and find the two to be similar, as I do here.

Table 7
Estimated Variance Components

	Var(α)	Var(λ)	Cov(α, λ)	θ	ρ	Var(ε)
(1)	0.131	—	—	—	0.892	0.061
	(0.010)				(0.008)	(0.003)
(2)	0.171	—	—	—	0.836	varies
	(0.008)				(0.010)	
(3)	0.266	0.000247	−0.0064	—	0.774	0.085
	(0.054)	(0.000044)	(0.0035)		(0.020)	(0.005)
(4)	0.351	0.000212	−0.0066	—	0.778	varies
	(0.050)	(0.000033)	(0.0029)		(0.012)	
(5)	0.158	—	—	−0.138	0.880	varies
	(0.008)			(0.021)	(0.007)	
(6)	0.449	0.000202	−0.0077	−0.121	0.860	varies
	(0.069)	(0.000041)	(0.0039)	(0.029)	(0.022)	

Note: Var(α) = variance of permanent component of income-to-needs ratio
Var(λ) = variance of individual-specific slope
ρ = $AR(1)$ parameter
θ = $MA(1)$ parameter
Var(ε) = variance of transitory component

rows of the table contain results from a model that contains an individual-specific slope, as well as an individual-specific intercept.[24] In this model Equation 15 is replaced with

(15a) $\varepsilon_{it} = \alpha_i + \lambda_i \text{age}_{it} + u_{it}.$

The variance of the individual-specific slope term is statistically significant in all cases. The estimates (from Row 3) of the variance of λ and the covariance of α and λ predict an increase in the cross-sectional variances of .14 from 1969 through 1988, capturing some of the increase in the empirical variances over time. Allowing the transitory component to vary across years has only minor effects on the other estimated parameters.

The final model estimated allows the transitory component of log income-to-needs to follow an ARMA(1,1) process, with a moving average parameter θ along with the autoregressive parameter ρ. The introduction of a moving average component is suggested by the work of Abowd and Card (1989), who find evidence of serially uncorrelated measurement error in earnings. The inclusion of a moving average component here will capture a purely transitory component of the variance structure, possibly due to measurement error in the income-to-needs ratio. These results show a small, negative estimated value for θ, consistent with the patterns in the empirical covariance matrix.

24. This type of model was originally proposed for earnings by Hause (1977), and has also been used by Lillard and Weiss (1979) and Baker (1997).

I next use the estimates in Table 7 to derive the expected time spent below the poverty line. To do this, I simulate the distribution of log income-to-needs, using the estimated variance components. I generate up to 20 years of data for individuals starting the period at age 30,[25] assuming normal distributions for v_{it} and α_i. In the simulations based on models including an individual-specific slope, I assume that α_i and λ_i follow a bivariate normal distribution with the estimated variances and covariance. These simulations hold business cycle conditions constant at the average of the calendar year-specific intercepts from the first-stage regressions (Equation 14). This generates a set of 20-year profiles of the income-to-needs ratio.

Because the hazard rate estimates refer to a flow-based distribution—the distribution of times poor for those just beginning a poverty spell—the predictions from the components-of-variance models must be presented in a similar form. The simulated data can be used to answer the same question as the hazard model—what is the distribution of total years poor over the next S years for individuals just starting a poverty spell? This is achieved by identifying each individual's first entrance into poverty during the simulated period. These individuals are then followed over the next ten years, and the resulting total times in poverty calculated. Simulations were run based on several of the models summarized in Table 7; the results discussed below are based on the results shown in Row 5, although those based on alternative specifications produced similar results.

As with the hazard models, I perform this exercise for eight demographic groups based on whether the household is headed by a female, the education of the household head, and race. Resulting distributions of time spent in poverty over the next ten years by persons just beginning a poverty spell are shown in Columns 2 and 5 of Table 8. For comparison, the table includes corresponding estimates from the hazard model and estimates from directly tabulating total time poor for a cohort of individuals observed entering poverty.[26] A complication for the "direct tabulation" results is that once additional covariates are introduced into the model, sample sizes for the direct tabulations are dramatically reduced. In order to increase sample sizes, the direct tabulation figures are based on all adults aged 20 to 55 entering poverty, with the given education and sex of the head. The results in Table 8 are for household heads with less than a high school education; in the interest of brevity results for those with more education are not shown, but the comparisons for the more educated groups are similar.

Table 8 shows that the hazard rate and direct tabulation results are generally consistent with one another, while the variance components model produces slightly different distributions. I focus first on the results for individuals beginning a poverty spell in a male-headed household (and remaining in a male-headed household over the next ten years). Among blacks all three distributions are similar, and result in a mean time in poverty over the next ten years of 4.2 to 4.3 years. In the results for whites, the variance components model predicts longer stays in poverty than the other two techniques. The average time in poverty is estimated at four years by this

25. An alternative would be to condition on the actual age distribution in the sample. I use a cohort beginning the period at a particular age for comparisons with the results of the hazard model.

26. The hazard model results here differ from those reported in the lower portion of Table 6 only by the starting age, which is 30 in this case.

Table 8
Comparison of Hazard Rate Estimates, Variance Components Estimates, and Direct Tabulations

	Male Head of Household			Female Head of Household		
Years Poor	Hazard Rate	Variance Components	Direct Tabulations	Hazard Rate	Variance Components	Direct Tabulations[a]
			Blacks			
1	17.3	20.6	18.5	3.3	14.8	8.7
2	16.0	14.8	23.8	4.7	11.9	4.0
3	13.5	12.3	10.1	5.8	11.0	10.0
4	11.5	10.6	8.1	7.0	10.5	4.4
5	9.2	9.1	5.2	8.5	9.5	2.9
6	8.1	8.0	14.3	9.6	8.9	6.3
7	7.1	7.0	1.7	10.5	8.3	11.3
8	6.3	6.1	5.5	12.6	7.7	8.8
9	4.5	5.5	9.0	14.7	7.6	15.9
10	6.5	6.0	3.8	23.2	9.8	27.9
Mean	4.4 years	4.30 years	4.2 years	7.0 years	5.0 years	6.9 years
N			307			146
			Whites			
1	32.9	24.1	39.5	13.3	17.7	20.4
2	21.4	16.2	16.2	11.6	13.7	18.5
3	13.8	12.9	9.3	10.6	12.1	14.3
4	9.8	10.7	14.0	10.4	10.5	4.6
5	6.6	8.9	1.5	9.8	9.5	7.8
6	4.4	7.3	6.6	8.5	8.4	4.6
7	3.2	6.1	2.9	7.9	7.7	5.6
8	2.5	5.3	2.5	7.8	6.8	9.1
9	2.5	4.2	5.4	7.8	6.4	12.2
10	2.9	4.5	2.1	12.3	7.1	2.8
Mean	3.1 years	4.0 years	3.1 years	5.2 years	4.6 years	4.3 years
N			217			71

Note: Distributions shown are simulated for individuals beginning a stay in poverty at age 30, and where the household head has less than a high school education.
a. For this category only, tabulations are based on individuals spending six or more of the ten years in a female-headed household. All other columns hold sex of the head constant over the full ten years.

method, compared to 3.1 years from the hazard model and direct tabulations. The variance components model also predicts substantially longer stays in poverty than the other two methods among more educated blacks and whites.

For individuals in households headed by females, the comparison across methods is less straightforward. Particularly for whites in female-headed households, the sample sizes available for the direct tabulations over ten years are too small ($N = 37$) to allow for useful inference. To increase sample sizes for the direct tabulations of white female-headed households I include individuals spending six or more out of ten years with a female head, rather than requiring female headship in all ten years. This leads to an expectation that the direct tabulation estimates for this group will result in less poverty persistence than the other two methods, which continue to simulate households with a female head in all ten years.

For black female heads the direct tabulation and hazard rate models are similar in terms of their predicted mean times in poverty of 6.9 and seven years out of the next ten. The two distributions are slightly different, with the direct tabulations showing a substantially larger proportion of individuals experiencing very short (three years of less) stays in poverty than the hazard rate. The variance component model, however, produces a distribution that is very different from both, and that suggests a mean time in poverty for this group of only five years. Among white female heads the hazard rate and direct tabulations differ substantially, although in a predictable way given that the direct tabulation results include women experiencing some of the ten years in a male-headed household, and so having shorter average stays in poverty.

Although these comparisons of the hazard and variance components models obviously do not constitute a formal test of the models' accuracy, the evidence in Table 8 suggests that the hazard model reproduces observed patterns of poverty persistence somewhat better than the variance components model. This does not seem to be sensitive to the particular version of the variance components model used for these simulations. The models summarized in Table 8 are a reasonable representation of those in the current literature on income dynamics, and none of these models suggests substantially different distributions of time in poverty.

One possible explanation for differences between the two approaches is that the variance components model estimated on a pooled sample of all adults is not flexible enough to capture different dynamic patterns across groups. In the current implementation, only differences in mean income-to-needs ratios are driving the differences across groups, since a common variance structure is estimated for the entire sample. To investigate this possibility, the variance components and hazard models were estimated for a sample restricted to white male heads. The results, and the comparison between the two models, were qualitatively the same as those discussed above.

Perhaps the most probable explanation for the differences in results based on the two estimation methods is that they are based on substantially different assumptions and impose different types of structure on the data. The components-of-variance approach is based on estimation of parameters describing the full distribution of income (or specifically the income-to-needs ratio) while the hazard model is based on samples of individuals with income-to-needs below the poverty line. It is thus not surprising that the hazard model better replicates observed patterns at the low end of the distribution. It is possible that income dynamics well below the mean of the distribution are different from those at or above the mean. Conventional variance components models, however, assume that the same dynamic patterns operate over all ranges of the income distribution.[27] For applications that focus on individuals in a particular portion of the income distribution, such as those below the poverty line, the hazard approach used here captures dynamics and total time in poverty well.

VII. Conclusion

Poverty persistence, as measured by the expected time poor for individuals starting a stay in poverty, is much greater than that predicted by previous work on single spells. Bane and Ellwood (1986) emphasized the substantial persistence of poverty that results from looking at the stock of persons poor at any point in time. They contrasted this with the lesser persistence in the flow measure—the expected time spent poor for an individual just beginning a poverty spell. My findings demonstrate that even flow-based distributions of time spent below the poverty line suggest substantial poverty persistence once reentry probabilities and multiple spells are considered.

Predicted times in poverty from the hazard model have also been compared with predictions from two other methods. Estimates from direct tabulations and from the hazard model are generally in agreement; across a variety of demographic groups the two methods yield similar distributions of years poor. Estimation of components-of-variance models of the income-to-needs ratio provides a third method for evaluating the persistence of poverty. This approach is quite common in the income dynamics literature, but predictions from these models are generally less consistent with observed patterns of poverty dynamics than are those from the hazard models.

All of the results presented here show that reliance on single-spell measures of poverty persistence significantly overstates the degree of mobility out of poverty. A thorough review of the poverty dynamics literature by Gottschalk, McLanahan, and Sandefur (1994, p. 90) highlights the observation that ''most low-income people, including most blacks, will be poor for less than two years.'' This is true in the sense that most will then experience at least a single year out of poverty, but looking at total time in poverty over the next decade changes conclusions about the longer-term persistence of poverty. Average time in poverty over the next ten years from Table 2 is over four years; more than half of all blacks and around one third of whites falling into poverty will spend five or more of the next ten years in poverty.

The answer to the question of whether poverty is a transitory or a permanent state largely depends on characteristics of individuals and their families. Individuals in

27. This does not mean that variance components models might not be modified to better capture poverty dynamics. The macroeconomics literature contains explorations of modifications to linear ARMA processes that better fit the dynamics of certain macro series. Neftci (1984) and Sichel (1993) offer evidence of asymmetries in macroeconomic series, and show that these asymmetries are missed by standard ARMA models. Hamilton (1989) builds a model in which a time-series process is subject to discrete regime shifts over time.

two-parent households experience the most transient poverty, with education and race also playing important roles in predicting stays below the poverty line. The average stay in poverty over the next ten years for those in households headed by black, less-educated males is approximately four years; for those in households headed by whites with at least a high school education the average stay is less than two and a half years. For individuals in households where the head is a single female, or has less than a high school education, poverty is an extremely persistent state. Among adults in female-headed households who fall below the poverty line, between 26 and 64 percent (depending on race and education level) will live below the poverty line for six or more of the next ten years. Among children in female-headed households the comparable figures range from 47 to nearly 90 percent. The conventional view that most individuals falling into poverty experience very short stays below the poverty line should be modified to account for the frequency and importance of multiple spells of poverty.

Appendix

Calculation of Standard Error Estimates Using Balanced Half-Sample Replications

This section describes the estimation of standard errors in Tables 1 through 5 using balanced half-sample replications; see Kish and Frankel (1970) or Wolter (1985), Chapter 3, for more details. This method makes use of variables on the PSID that assign each household to one of 32 different strata, and to one of two primary sampling units within each stratum.

Given these assignments, a series of half-samples is created by selecting one of the two sampling units within each stratum. A large number of different half-samples is possible, given the various possible selections across all 32 strata. In order to minimize the variance of the variance estimator, the half-samples are selected in a particular way. Wolter (1985) provides a series of orthogonal matrices (known as

Table A1
Sample Sizes

	Blacks		Whites	
	Exit Hazard	Reentry Hazard	Exit Hazard	Reentry Hazard
1969–88 sample				
Persons	6,854	6,400	4,000	3,794
Person-years	36,783	44,826	13,723	31,760
1973–88 sample				
Persons	7,157	6,040	3,662	3,653
Persons-years	42,902	39,016	12,910	28,995

Table A2
Sample Means (over all person-years), Spells In and Out of Poverty, 1969–88

	Blacks		Whites	
	In Poverty	Out of Poverty	In Poverty	Out of Poverty
Household income ($1987)	$8,721	$29,589	$7,841	$31,957
Income to needs ratio	0.60	2.21	0.65	2.64
Household size	4.32	3.84	3.40	3.35
Fraction with head's education less than thigh school	0.71	0.62	0.69	0.54
Fraction with single female head	0.63	0.28	0.42	0.20
Age of household heads	41.8	41.8	46.2	41.6
Age of all individuals	25.9	31.7	35.9	34.2

Hadamard matrices) that describe which unit should be selected from each of the strata to guarantee that the variance estimator is as precise as possible.

Let $\hat{\beta}$ be a parameter estimate based on the entire sample. Then, from each of 32 different half-samples, let $\hat{\beta}_\alpha$ be the same parameter estimated from the αth half-sample. The variance estimator, which can be shown to be a consistent estimator under certain regularity conditions (Wolter 1985, Appendix B), is then given by:

$$v_{bhr}(\hat{\beta}) = \sum_{\alpha=1}^{32} (\hat{\beta}_\alpha - \hat{\beta})^2/32.$$

In the case of a linear estimator, this formula, given the strategic selection of half-samples described above, will reproduce exactly the "textbook" variance estimator for a complex survey design. In the case of non-linear estimators the finite-sample correspondence is approximate. Wolter (1985) describes studies establishing the appropriateness of balanced half-sample variance estimators in a variety of specific nonlinear applications.

Table A3
Estimated Covariance Matrix

	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88
69	**.301**	.772	.697	.632	.639	.624	.585	.564	.560	.507	.516	.460	.434	.429	.410	.385	.328	.327	.327	.323
70	.234	**.315**	.832	.770	.754	.654	.630	.624	.589	.531	.529	.503	.467	.448	.464	.457	.405	.401	.393	.376
71	.232	.278	**.367**	.786	.769	.687	.638	.621	.592	.562	.538	.518	.479	.485	.481	.457	.411	.408	.397	.387
72	.203	.249	.279	**.343**	.849	.732	.701	.666	.650	.593	.557	.546	.542	.528	.505	.481	.397	.411	.434	.424
73	.199	.236	.264	.282	**.321**	.791	.724	.708	.686	.646	.606	.584	.554	.530	.497	.467	.412	.417	.420	.425
74	.204	.215	.248	.256	.267	**.356**	.787	.775	.734	.699	.647	.631	.615	.594	.567	.532	.446	.439	.442	.433
75	.188	.203	.226	.240	.240	.275	**.342**	.888	.787	.741	.674	.644	.656	.612	.589	.526	.479	.480	.458	.452
76	.180	.200	.219	.227	.233	.269	.302	**.376**	.877	.805	.729	.696	.698	.643	.606	.561	.508	.487	.465	.456
77	.184	.195	.215	.228	.233	.263	.276	.306	**.360**	.837	.761	.720	.716	.672	.644	.592	.496	.511	.509	.471
78	.166	.175	.203	.207	.219	.249	.259	.280	.300	**.357**	.818	.761	.725	.672	.648	.590	.519	.515	.505	.473
79	.169	.174	.194	.194	.205	.230	.235	.253	.272	.291	**.355**	.817	.754	.718	.705	.635	.600	.586	.569	.525
80	.157	.173	.195	.199	.205	.234	.234	.252	.269	.283	.302	**.386**	.821	.728	.687	.606	.549	.551	.510	.529
81	.163	.177	.199	.218	.215	.252	.263	.278	.295	.297	.308	.350	**.470**	.787	.745	.683	.593	.587	.571	.539
82	.170	.179	.212	.223	.217	.256	.259	.270	.291	.290	.309	.327	.390	**.522**	.793	.687	.616	.621	.631	.579
83	.160	.183	.207	.211	.201	.241	.245	.251	.275	.276	.299	.304	.364	.408	**.507**	.786	.735	.703	.669	.580
84	.157	.188	.206	.210	.197	.236	.229	.243	.264	.262	.281	.280	.348	.369	.416	**.553**	.780	.718	.690	.599
85	.139	.172	.191	.179	.180	.205	.216	.228	.229	.239	.275	.263	.313	.343	.403	.446	**.592**	.795	.710	.614
86	.126	.155	.173	.169	.165	.183	.197	.198	.215	.216	.244	.240	.282	.314	.351	.374	.428	**.491**	.808	.681
87	.135	.163	.180	.191	.179	.198	.201	.203	.229	.226	.255	.238	.294	.343	.358	.385	.410	.425	**.564**	.784
88	.123	.144	.163	.172	.167	.179	.183	.184	.196	.196	.217	.228	.256	.290	.286	.309	.327	.331	.408	**.481**

Note: Below the diagonal entries are covariances between indicated years; above the diagonal entries are correlation coefficients.

References

Abowd, John, and David Card. 1989. "On the Covariance Structure of Earnings and Hours Changes." *Econometrica* 57(2):411–45.

Baker, Michael. 1997. "Growth Rate Heterogeneity and the Covariance Structure of Earnings." *Journal of Labor Economics* 15(2):338–75.

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

Cutler, David. 1995. "The Incidence of Adverse Medical Outcomes Under Prospective Payment." *Econometrica* 63(1):29–50.

Duncan, Greg. 1983. "The Implications of Changing Family Composition for the Dynamic Analysis of Family Economic Well-Being." In *Panel Data on Incomes,* ed. A. B. Atkinson, 203–39. London: London School of Economics.

Duncan, Greg, et al. 1984. *Years of Poverty, Years of Plenty.* Ann Arbor: Survey Research Center.

Duncan, Greg, and Willard Rodgers. 1991. "Has Children's Poverty Become More Persistent?" *American Sociological Review* 56(4):538–50.

Eberwein, Curtis, John Ham, and Robert LaLonde. 1997. "The Impact of Being Offered and Receiving Classroom Training on the Employment Histories of Disadvantaged Women: Evidence from Experimental Data." *Review of Economic Studies* 64(4):655–82.

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

Gottschalk, Peter. 1982. "Earnings Mobility: Permanent Change or Transitory Fluctuations?" *Review of Economics and Statistics* 64(3):450–56.

Gottschalk, Peter, Sara McLanahan, and Gary Sandefur (1994). "The Dynamics and Intergenerational Transmission of Poverty and Welfare Participation." *In Confronting Poverty: Prescriptions for Change,* ed. Sheldon Danziger, Gary Sandefur, and Daniel Weinberg, 85–108. Cambridge: Harvard University Press.

Gottschalk, Peter, and Robert Moffitt. 1994. "Welfare Dependence: Concepts, Measures, and Trends." *American Economic Review* 84(2):38–42.

Ham, John, and Samuel Rea. 1987. "Unemployment Insurance and Male Unemployment Duration in Canada." *Journal of Labor Economics* 5(3):325–53.

Hamilton, James. 1989. "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle." *Econometrica* 57(2):357–84.

Hause, John. 1977. "The Covariance Structure of Earnings and the On-the-Job Training Hypothesis." *Annals of Economic and Social Measurement* 6(4):335–65.

Heckman, James, and Burton Singer. 1984. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica* 52(2):271–320.

Heckman, James, and James Walker. 1990. "The Relationship Between Wages and Income and the Timing and Spacing of Births: Evidence from Swedish Longitudinal Data." *Econometrica* 58(6):1411–41.

Hill, Martha 1981. "Some Illustrative Design Effects: Proper Sampling Errors Versus Simple Random Sample Assumptions." In *Five Thousand American Families—Patterns of Economic Progress.* Volume IX, ed. Martha Hill, Daniel Hill, and James Morgan, 457–66. Ann Arbor: Institute for Social Research.

Kish, Leslie, and Martin Frankel. 1970. "Balanced Repeated Replication for Standard Errors." *Journal of the American Statistical Association* 65(331):1071–94.

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

Lillard, Lee, and Robert Willis. 1978. ''Dynamic Aspects of Earnings Mobility.'' *Econometrica* 46(5):985–1012.

Moffitt, Robert, and Michael Rendall. 1995. ''Cohort Trends in the Lifetime Distribution of Female Family Headship in the U.S., 1968–1985.'' *Demography* 32(4):407–24.

Neftci, Salih N. 1984. ''Are Economic Time Series Asymmetric over the Business Cycle?'' *Journal of Political Economy* 92(2):307–28.

Sichel, Daniel. 1993. ''Business Cycle Asymmetry: A Deeper Look.'' *Economic Inquiry* 31(2):224–36.

Stevens, Ann Huff. 1994. ''The Dynamics of Poverty Spells: Updating Bane and Ellwood.'' *American Economic Review* 84(2):34–37.

———. 1995. ''Essays on Income Dynamics: Long-Term Effects of Job Displacement and Measuring the Persistence of Poverty Over Multiple Spells.'' Ph.D. Dissertation, University of Michigan.

Wolter, Kirk. 1985. *Introduction to Variance Estimation.* New York: Springer-Verlag.