DYNAMIC ASPECTS OF EARNING MOBILITY[1]

by

Lee A. Lillard
National Bureau of Economic Research

and

Robert J. Willis
National Bureau of Economic Research
and
Stanford University

Revision of Working Paper No. 150, Sept 76

April, 1977
(revised)*

Abstract

This paper proposes an econometric methodology to deal with life cycle earnings and mobility among discrete earnings classes. First, we use panel data on male log earnings to estimate an earnings function with permanent and serially correlated transitory components due to both measured and unmeasured variables. Assuming that the error components are normally distributed, we develop statements for the probability that an individual's earnings will fall into a particular but arbitrary time sequence of poverty states. Using these statements, we illustrate the implications of our earnings model for poverty dynamics and compare our approach to Markov chain models of income mobility.

*This draft supersedes an earlier version of this paper which was distributed as NBER Working Paper No. 150, September 1976.

1. INTRODUCTION

Is poverty a transitory status or a permanent condition of individuals and households? More broadly, is there a high or low degree of mobility over time in an individual's place in the distribution of earnings or a household's place in the distribution of income? The increasing availability of longitudinal data files containing information on individual and family earnings, income, personal characteristics and environmental variables together with a growing body of theoretical literature on life cycle behavior have greatly enhanced the capacity of economists to answer such questions. The purpose of this paper is to propose an econometric methodology that may serve as a link between theory and data on the dynamic aspects of earnings and of income distributions. In the case of earnings mobility -- the case considered in this paper -- it has the added advantage of providing a direct linkage with traditional human capital earnings functions.

Although human capital earnings functions are cast in a life cycle framework, it is only rarely that the availability of longitudinal data has been deemed crucial for their estimation. The reason for this can probably be traced to the emphasis in these models on the paradigm of the "representative man." This paradigm leads to an empirical emphasis on the effects of independent variables on the mean earnings of an average individual and to a relative lack of interest in individual differences.[2] One consequence of this is that most human capital earnings functions are incapable of describing the life cycle dynamics of the earnings distribution; they simply describe the average growth path of earnings for the representative individual.[3]

In contrast, longitudinal data have played an essential role in several recent studies by economists of mobility among discrete income or earnings classes. (McCall [1973], Levy [1975], Schiller [1976] Schorrocks [1976]). A central question in these studies is the extent to which an individual retains his position in the earnings or income distribution of his cohort over his life cycle. From this point of view, cross-sectional income distributions are susceptible to a wide variety of interpretations. One extreme possibility is complete income stratification. In this case, knowledge of an individual's position in the income distribution in a cross-sectional survey in year t is a perfect predictor of his position in subsequent years. Levy [1976], for example, argues that this interpretation of cross-section statistics on the poverty population was implicit in the view of government poverty policy-makers during the 1960's: Without outside help, an individual who is in poverty cannot expect to get out of poverty. As Levy points points out, this possibility cannot be distinguished in cross-section data from the possibility that poverty is merely a transient status in which many of the individuals who are in poverty in one year will be replaced by others who were initially out of poverty. Thus, at the opposite extreme from complete income stratification is complete income mobility in which an individual's probability of being in some discrete income class (e.g., poverty or nonpoverty, a given decile of the earnings distribution, etc.) in a given period is independent of his prior income status. While neither of these extreme possibilities seems realistic, it is clear that longitudinal data on income or earnings are necessary to establish what reality is.

In this paper, we propose a fairly simple methodology to deal with life cycle earnings and mobility among discrete earnings classes. First, we estimate an earnings function with log male earnings as the dependent variable, using seven years of earnings data from the University of Michigan Income Dynamics Panel, 1967-73. We estimate permanent and serially correlated transitory components of earnings due to both measured and unmeasured variables. Measured variables are represented by an earnings function, while unmeasured variables are represented by components of residual variance. Assuming that the error components are normally distributed, the intertemporal distribution of log earnings, conditional on measured variables, is multivariate normal with a correlation matrix determined by the estimated components. The probability that an individual's earnings will fall into a particular, but arbitrary time sequence of discrete earnings classes is then computed by evaluating the appropriate multivariate normal integral within the limits given by the definition of the earnings classes.[4]

It is important to point out that our approach to earnings mobility copes quite easily with certain issues that have proven difficult in Markov chain mobility models such as those of Champernowne (1953), McCall (1973) and Schorrocks (1976) which provide the major alternative methodology. First in a Markov model it is necessary to define in advance the number and width of the income classes among which transitions take place. While, for example, Champernowne's theory assumes an infinite number of classes of equal logarithmic length, data limitations necessitate a relatively small number of classes to preserve cell size. In addition, policy interests may suggest focussing on certain classes such as "poverty" because of their normative rather than their behavioral significance. This presents a difficulty because a process which is Markovian in a given state space need not be Markovian when states are redefined (see Shorrocks, 1976). In our approach, the number and width of income classes may be defined arbitrarily. Second, Markov models present some difficult and subtle statistical problems when transition probabilities vary among observationally identical individuals (i.e. individuals with identical measured characteristics).[5] Because we estimate the components due to unmeasured permanent and serially correlated components, we can deal with this problem of population heterogeneity by computing directly the distribution of probabilities of an arbitrary sequence of earnings states. Finally, our approach allows transition probabilities to vary across people and over time because of variations in measured variables or because of variations in the definition of earnings classes. Thus, there is no need to assume that transition probabilities remain constant over time, as is often done in the Markov models (e.g. Shorrocks (1976)).

The organization of the paper is as follows. In Section 2, we specify an earnings function that utilizes all sample information about earnings and is clearly linked to the life cycle earnings literature. By considering both permanent and serially correlated transitory components as well as measured variables that vary across time and people, we take fairly full account of population heterogeneity. To some extent, serial correlation also captures some of the effects of the arbitrary time frame of the data (i.e. yearly earnings). In Section 3, we analyze the implications of this earnings function for earnings mobility across discrete earnings classes. For simplicity, we develop probability statements for two earnings classes. These probabilities are illustrated empirically in Section 4 for transitions into and out of poverty by blacks and whites. The state "poverty" is arbitrarily defined by a poverty line equal to one-half the median of U.S. male earnings in each year. For expositional convenience empirical probability statements used to illustrate the general methodology are based on a variance component model which omits measured variables.

2. EMPIRICAL EARNINGS MODEL

The basic ingredient of our approach to earnings mobility is an empirical earnings function. The purpose of this section is first to outline the specification of the earnings function, its error structure and the statistical methodology used to estimate it and, second, to present parameter estimates based on seven years of earnings data from the Income Dynamics panel.

The earnings function is of the form

$$(2.1) \qquad Y_{it} = \underset{\sim}{X}_{it}\beta + \Gamma_t + \mu_{it} \qquad i=1,\ldots,N; \quad t=1,\ldots,T$$

where Y_{it} is the natural logarithm of the real annual earnings (in 1970 dollars) of the ith person in the t-th year. Each person is observed each of the same T years. Three successively more complex forms of this earnings function will be estimated: the models including (1) Γ_t (i.e., time dummies) only, (2) a conventional earnings function, $\underset{\sim}{X}\underset{\sim}{\beta}$, including race, education, labor force experience and time effects and (3) a fairly comprehensive earnings function, $\underset{\sim}{X}\underset{\sim}{\beta}$, which includes such additional variables as detailed job histories, geographical data, an index of local labor market conditions, union membership and interactions of certain variables with time. (A full list of these variables is given in the Appendix.)

The error structure is assumed to be of the following form.

$$(2.2) \qquad \mu_{it} = \delta_i + v_{it}$$

and

$$v_{it} = \gamma v_{it-1} + \eta_{it}$$

where δ_i is a random individual component ($\sim (0,\sigma_\delta^2)$), η_{it} is a purely random component ($\sim_{iid}(0,\sigma_\eta^2)$) and γ is the serial correlation coefficient common to all individuals. The variates δ_i, η_{it} are assumed to be independent of each other and of $\underset{\sim}{X}_{it}$ and Γ_t. This error structure, which we call the "autocorrelated individual component model," combines the features of both the traditional variance component model of individual random effects and first order autocorrelation over time net of the individual component.[6]

The individual component of this error structure, δ, represents the effect of unmeasured individual variables in equation (2.1). It may be termed population heterogeneity in mean relative earnings and will play a central role in the analysis of the dynamics of poverty. In this section it will be assumed that σ_δ^2 is homogenous throughout the population. Later we will illustrate heterogeneity over identifiable subgroups of the population with respect to σ_δ^2, σ_η^2 and γ.

The serial correlation term, γ, may be interpreted in a couple of ways. First, it reflects the effect of random shocks which persist longer than one year but which deteriorate in effect over time. Second, it reflects the operation of individual, unobserved variables which are serially correlated over time, i.e., change slowly through time.

Time effects are estimated as fixed parameters (i.e., dummy variables). While in equation (2.1) time effects are shown for simplicity to be the same for all individuals, they are different for various identifiable population subgroups (e.g., race). Year effects capture the combined effect of time-varying macroeconomic variables such as exogenous productivity changes, market conditions, etc., that are not explicitly entered as exogenous variables in our model. In these data time effects are estimated directly, rather than used as a random component, since there are only seven years in the panel and more than six parameters are estimated leaving far too few degrees of freedom for the latter procedure to be interpretable (See Nerlove [1971]).

Assuming that initial earnings at the beginning of the work history are shocked by an error of the form

$$(2.3) \qquad v_{i1} = \eta_{i1}/\sqrt{1-\gamma^2}$$

and η_{it} thereafter[7] $(t > 1)$ the residual covariance structure is of the form

$$(2.4) \qquad E(\mu_{it}\mu_{j\tau}) = \begin{cases} \sigma_\delta^2 + \sigma_v^2 = \sigma_\mu^2 & i=j,\ t=\tau \\ \sigma_\delta^2 + \gamma^S\sigma_v^2 = \sigma_\mu^2[\rho + (1-\rho)\gamma^S] & i=j,\ |t-\tau| = S > 0 \\ 0 & i\neq j \end{cases}$$

where

$$(2.5) \qquad \sigma_v^2 = \sigma_\eta^2/(1-\gamma^2)$$

and

$$\rho = \sigma_\delta^2/\sigma_\mu^2 \,. \tag{2.6}$$

Correspondingly, the residual covariance matrix for each individual (δ given) is

$$\underset{\sim}{\Sigma}^* = \sigma_v^2 \begin{bmatrix} 1 & \gamma & \gamma^2 & \cdots & \gamma^{T-1} \\ \gamma & 1 & \gamma & \cdots & \cdot \\ \gamma^2 & \gamma & 1 & \cdots & \cdot \\ \cdot & \cdot & \cdot & \ddots & \\ \cdot & \cdot & \cdot & & \cdot \\ \gamma^{T-1} & \cdot & \cdots & \cdots & 1 \end{bmatrix} \tag{2.7}$$

and the aggregate covariance matrix (over individuals or for a random individual) is

$$\underset{\sim}{\Sigma} = \underset{\sim}{\Sigma}^* + \sigma_\delta^2 \underset{\sim}{\iota}\underset{\sim}{\iota}' \tag{2.8}$$

where $\underset{\sim}{\iota}$ is a T x 1 vector of "1"'s/

The Data

The University of Michigan Panel Study of Income Dynamics represents a panel of seven contiguous years, 1967-73. The survey included 5,517 households, about 2,000 from the 1967 Survey of Economic Opportunity and an additional 3,000 from a cross section of U.S. families. Since the SEO subsample selected by the Survey Research Center for inclusion in the panel was not random with respect to income this group is excluded from consideration.[8] A detailed description of the data is provided in Morgan (1974). The parameters of the model are estimated for the 1,041 white and 103 black persons identified as a male head of household (including single persons) between the ages 18 and 58 in 1967 who were not disabled, retired or a fulltime student during the period and who reported positive annual hours and earnings each year. Means and standard deviations of variables are presented in Appendix Table A1. All earnings values are in real 1970 dollars.

Parameter Estimates

The parameters $\underset{\sim}{\beta}$ and Γ_t are estimated by OLS on the data pooled over individuals and years.[9] The parameters σ_δ^2, σ_η^2 and γ are estimated by maximum likelihood from the OLS residuals.[10] Parameter estimates for the residual structure are reported in Table 1 for alternative sets of independent variables including (1) no independent variables, except Γ_t, (2) a simple set of independent variables including only schooling, experience, race and Γ_t, and (3) a more comprehensive set of independent variables including individual and job related characteristics, some indicators of local market conditions, time and interactions.[11]

First consider the components of variation in log earnings controlling only for individual year differences, i.e., $Y_{it} = \Gamma_t + \mu_{it}$. Since the variance in the log of earnings is often used as an index of earnings inequality, these components may be thought of as sources of earnings inequality. Total within year variance is .307. Permanent earnings differences among individuals represent 73.1 percent of total variation. Of the 26.9 percent remaining stochastic variation from period to period, 22.4 percentage points may be considered purely stochastic variation, the remainder being due to serial correlation. The permanent component may be interpreted as the effect of permanent differences among individuals, some of which we observe and others that we do not. As noted earlier, serial correlation may represent the effect of serially correlated independent variables or the effect of transitory variables whose effects last more than a year. We will observe some of these variables but not all. The

Table 1. Components of Variance, Autocorrelated Individual Component Model

	No. Obs.	$\hat{\sigma}^2_\mu$	$\hat{\sigma}^2_\delta$	$\hat{\sigma}^2_\eta$	$\hat{\sigma}^2_v$	$\hat{\gamma}$	$\hat{\rho}$
Log Earnings							
All	1144	.307	.224	.069	.083	.406	.731
Blacks	103	.369	.299	.059	.070	.395	.811
Whites	1041	.291	.207	.070	.084	.408	.711
Residual-Simple Eqn.*							
All	1144	.206	.125	.068	.081	.402	.606
Blacks	103	.219	.146	.060	.073	.419	.667
Whites	1041	.206	.124	.069	.082	.399	.602
Residual-Comp. Eqn.**							
All	1144	.153	.072	.071	.081	.350	.471
Blacks	103	.154	.081	.064	.073	.350	.526
Whites	1041	.153	.071	.072	.082	.350	.464

NOTE:

*Including only race, years of schooling, experience and experience squared and time dummies. The regression is in Appendix Table A2.

**Including a comprehensive set of explanatory variables. The regression is presented in Appendix Table A2.

purely stochastic component includes both the effect of transitory variables and measurement error. These two have quite different implications for poverty analysis but cannot be distinquished here.

Using estimates from the full model ($\hat{\rho}$ = .731, $\hat{\gamma}$ = .406) the correlation of earnings in adjacent years across individuals is $\hat{\rho} + (1-\hat{\rho})\hat{\gamma} = .840$. It declines to .775 for observations two years apart, .74.9 for three years apart, .738 for four years and, asymtotically to .731. These values are represented graphically by the solid line in Figure 1. Corresponding correlations from the simple variance component model ($\hat{\rho}$ = .790, $\gamma \equiv 0$) are represented by the horizontal dotted line in Figure 1 and correlations from the simple serial correlation model ($\rho \equiv 0$, $\hat{\gamma}$ = .841) by the negatively sloped dotted line. The superiority of the full model (i.e., autocorrelated individual component model) is clearly illustrated by its close fit to the actual pooled correlation denoted by dots in Figure 1. A maximum likelihood ratio test rejects each of the simple restrictions $\gamma \equiv 0$ and $\sigma_{\delta}^{2} \equiv 0$.

Our results indicate larger permanent and smaller transitory variances in log earnings among blacks compared to whites. The black permanent component is 44 percent larger than the white value. Much of this racial differential in permanent components is due to measured variables. Thus, controlling for a comprehensive set of explanatory variables, the unmeasured permanent component is only 14 percent larger for blacks. In contrast, the transitory component is about 16 percent smaller among blacks compared to whites. Total earnings variance is about 27 percent larger for blacks, but variance due to unmeasured factors (both permanent and transitory) is almost identical for blacks and whites when the comprehensive set of explanatory variables are held constant.

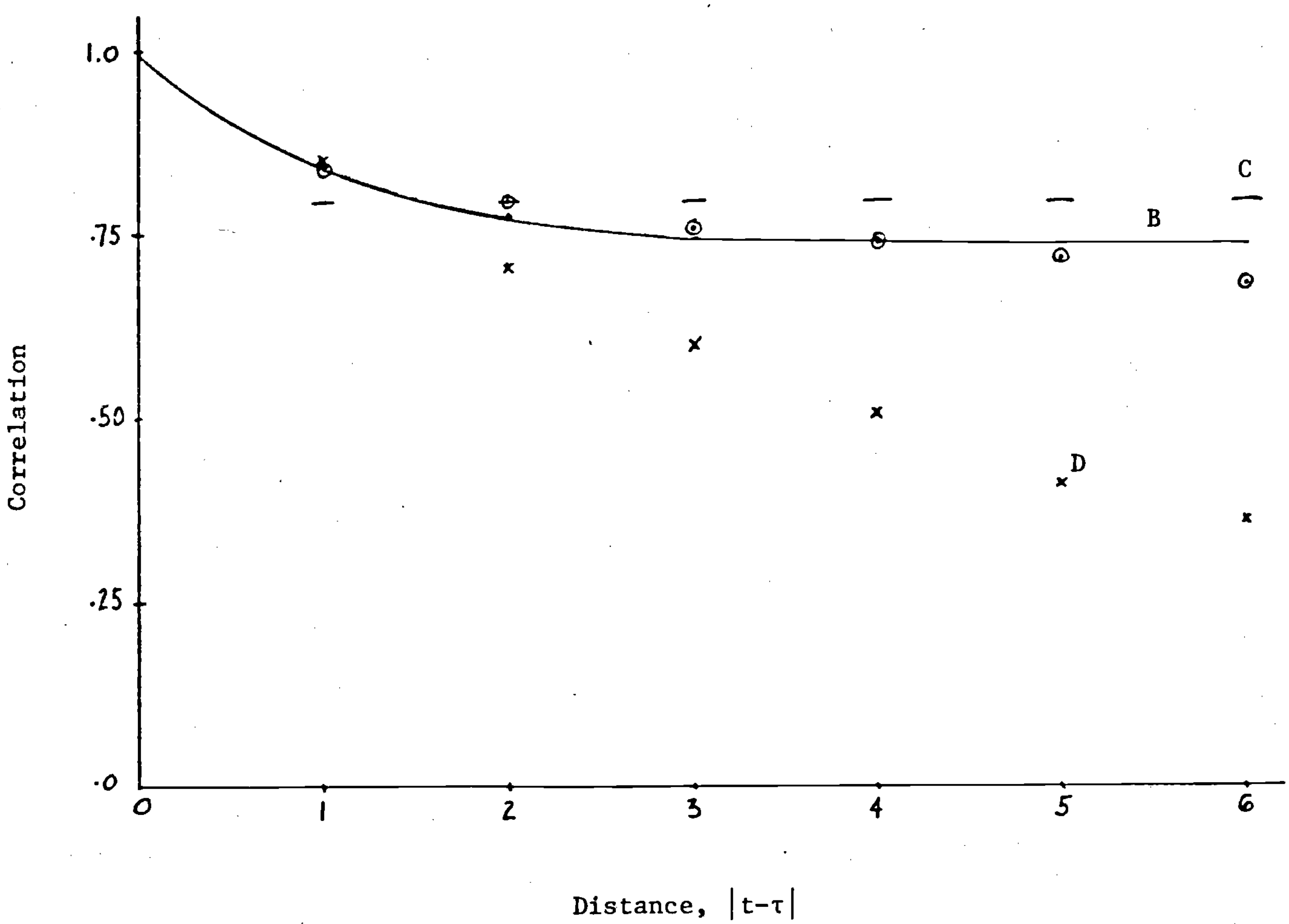

Figure 1. Actual and Predicted Correlations Among Years for Log Earnings

Legend:

Model		Symbol	Parameters ρ	Parameters γ
A	Actual	⊙	data	
B	Full Model	———	.731	.406
C	Var. Comp.	— —	.790	0
D	Serial Corr.	x x	0	.841

Consider the role of measured variables in these components of earnings variation.[12] First consider simply introducing race, years of schooling, and years of work experience into the earnings function. The earnings function is then

$$(2.9)\quad Y_{it} = 7.665 + .084\ \text{Sch} + .038\ \text{Exp} - .007\ \text{Exp}^2 - .166\ \text{Black} + \Gamma_t + \mu_{it}$$

The β parameter estimates conform to traditional estimates of many earlier studies. These variables explain 33 percent of total earnings variation but explain 44 percent of the permanent component.[13] As expected, these individualized variables explain none of the stochastic variation. This fairly simple earnings function then does quite well at explaining permanent earnings differences even though its performance would appear somewhat poorer in cross-sectional estimates in which the permanent and transitory components cannot be separated. Within racial groups, schooling and experience alone explain 51 percent of the permanent component for blacks and 40 percent for whites. The corresponding figures for total earnings variation are 41 percent and 29 percent respectively.

The overall explanatory power of the earnings function is increased to 50 percent when a fairly comprehensive set of measured variables are included. The full list of variables and estimated coefficients are presented in an appendix available from the authors. They include detailed job history data, an index of local labor market conditions, union membership, and interactions of certain variables with time to allow differential growth rates. A detailed discussion of these parameter estimates would be interesting but is not appropriate to the focus of this paper. Within racial groups the explanatory power of the X variables and Γ_t is 58 percent for blacks and 47 percent for whites. This is a consequence of the larger share of total variation due to permanent differences for blacks and the greater explanatory power of these variables with respect to the permanent component. The measured variables explain 46 percent of total residual variation for whites and 53 percent for blacks. Of total permanent variance, 73 percent is due to measured variables for blacks and only 65 percent for whites. The unmeasured permanent component accounts for 24 percent of total earnings variation for whites, 22 percent for blacks and 24 percent (including race) for the total population. Correspondingly, 71 percent of total earnings variation is accounted for by measured variables plus permanent differences among whites, 81 percent for blacks and 73 percent (including race) for the total sample. This set of variables, some of which vary a bit with time, only slightly affect the estimate of purely stochastic variance, σ_η^2 . A rather weak aspect of even the fairly comprehensive set of variables used here is that too few variables having transitory effects are observed. However, the estimate of serial correlation is reduced from about .40 to .35 for each group by this set of variables.

3. DYNAMIC ASPECTS OF POVERTY

In this section we develop mathematical expressions, based on the earnings function and the components of variance, for the probability of observing a given individual or group of individuals in any arbitrary sequence of poverty states. To be in a state of poverty at a given time, an individual's earnings must fall below some prespecified but arbitrary level. We also derive formulas for the joint and conditional probabilities of poverty status in two or more years and for the corresponding distributions of such probabilities over individuals. The conditional probabilities are compared to analogous probabilities in Markov chain models of poverty dynamics. The empirical counterparts of these formulas will be examined in Section 4. Although we assume that η and δ are normally distributed, our approach could be used to derive probability statements under less restrictive assumptions or assuming alternative probability distributions for δ and η.

The poverty level and the earnings class it defines are prespecified in the sense of being based on considerations outside the probability model itself. Thus, this level, denoted as Y^*_{it}, may vary from period to period and across well-defined groups of individuals (e.g., by place of residence or family size). And since the probability of being in poverty is invariant over monotonic transformations of earnings, we may use the logarithm of earnings directly.

Poverty Probabilities for an Individual

The probability of an individual being in poverty in year t, ϕ_{it}, is simply the probability that his earnings fall below the poverty line, Y^*_{it}. Using the notation of Section 2,

$$(3.1) \qquad Y_{it} \leq Y^*_{it} \iff \frac{v_{it}}{\sigma_v} \leq \frac{(Y^*_{it} - X_{it}\beta - \Gamma_t - \delta_i)}{\sigma_v} \equiv b^*_{it} .$$

Assuming normality for η and thus v

$$(3.2) \qquad \phi_{it} = F(b^*_{it})$$

where F denotes the cumulative standard normal function. Clearly, ϕ_{it} is a monotonic function of δ.

Consider the joint probability of poverty (not-poverty) in an arbitrary pair of years t and τ. Denote the joint probability of the four pairs of outcomes as follows: (1) the probability of being in poverty in both years t and τ by $\phi_{t,\tau}$; (2) the probability of being out of poverty both years $\phi_{\sim t,\sim\tau}$; and so forth. Clearly,

$$(3.3) \qquad \phi_{it,\tau} = F(b^*_{it},\ b^*_{i\tau};\ \gamma^{|t-\tau|}) \ ,$$

$$(3.4) \qquad \phi_{it,\sim\tau} = F(b^*_{it},\ -b^*_{i\tau};\ -\gamma^{|t-\tau|})$$

and so forth, where F is the cumulative standardized ($\sigma_1 = \sigma_2 = 1,\ \mu_1 = \mu_2 = 0$) bivariate normal with correlation $\gamma^{|t-\tau|}$.

The individual probability statements straightforwardly generalize to a sequence of poverty (not poverty) states for K arbitrary years. Let

$$(3.5) \qquad H_t = \begin{cases} 0 & \text{if in poverty in year } t \\ 1 & \text{if not in poverty in year } t \end{cases}$$

$$(3.6) \qquad J_t = \begin{cases} 1 & \text{if in poverty in year } t \\ -1 & \text{if not in poverty in year } t \end{cases}$$

and $\underset{\sim}{J}$ be a $K \times 1$ vector of values of J_t so that $\phi_{\underset{\sim}{H}}$ represents the probability of the sequence of states given by $\underset{\sim}{H}$. For any K contiguous years

$$\phi_{\underset{\sim}{H}} = F((-1)^{H_1} b^*_{it_1}, (-1)^{H_2} b^*_{it_2}, \ldots, (-1)^{H_k} b^*_{it_k} ; \frac{1}{\sigma_v^2} \underset{\sim}{J}'\underset{\sim}{\Sigma}^*\underset{\sim}{J}) \tag{3.7}$$

where $\underset{\sim}{\Sigma}^*$ is defined in equation (2.7).

The Distribution of Poverty Probabilities Among Individuals

It is important to distinguish the probability statements for individuals in the previous section from probability statements for groups of observationally identical individuals (i.e., individuals who share identical values of X_t). The individual probability statements are conditional on a given value of δ while δ varies across observationally identical individuals. This generates a distribution of poverty probabilities among such individuals which depends on the distribution of δ.

The distribution function of individual poverty probabilities ϕ_{it}, for observationally identical persons, assuming normality for δ, is given by

$$G(\pi_t) = F(\frac{\sigma_v}{\sigma_\delta} F^{-1}(\pi_t) - \frac{\sigma_\mu}{\sigma_\delta} b_{it}) \tag{3.8}$$ [14] .

This distribution function provides the answer to the question "What proportion of the population has less than a π_t probability of falling into poverty in any single year?". The corresponding density function, is

$$g(\pi_t) = \frac{\sigma_v \cdot f(\frac{\sigma_v}{\sigma_\delta} F^{-1}(\pi_t) - \frac{\sigma_\mu}{\sigma_\delta} b_{it})}{\sigma_\delta \cdot f(F^{-1}(\pi_t))} \tag{3.9}$$ [15]

where f denotes the standard normal density function. Analogous statements for the distributions of multiple period sequences (e.g., $\phi_{it,\tau}$ or $\phi_{it,\sim\tau}$)

can be derived. Some of these statements are reported in an earlier draft of this paper.

Aggregate Poverty Probabilities

The mean of the distribution of single year poverty probabilities corresponds to the probability that a random person in the group will be observed in poverty and to the expected proportion of the group falling into poverty. Note, however, that this probability is not generally equal to the single year poverty probability of the "representative" or "median" individual (with $\delta = 0$). Similarly, the probability of any given sequence of poverty states for the representative individual is not equal to the probability of that sequence for an individual chosen at random or to the expected proportion of the population who follow that sequence, because of population heterogeneity caused by variation in δ.

Since in the aggregate, allowing δ to vary,

$$(3.10) \qquad Y_{it} \leq Y^*_{it} \quad \Longleftrightarrow \quad \frac{\mu_{it}}{\sigma_\mu} \leq \frac{(Y^*_{it} - \underset{\sim}{X}_{it}\underset{\sim}{\beta} - \Gamma_t)}{\sigma_\mu} \equiv b_{it} ,$$

the aggregate probability of poverty in a single year t, P_t, is the probit function

$$(3.11) \qquad P_t = F(b_{it}) \quad .$$ [16]

In a single year, or for pooled independent cross sections where each individual is observed only once, the simple probit function is quite appropriate. Independent variables enter through b_{it} but probability statements can be applied only to populations and not to individuals. One cannot from this probability statement alone identify truth between the extremes "P_t percent of the population will always be in poverty and the remainder never" and "each individual faces a P_t probability of falling into poverty." The proportion of the population whose expected income (including δ) is below the poverty line is given by

$$(3.12) \qquad P_t' = \Pr(\delta \le b_{it} \cdot \sigma_\mu) = F(b_{it} \frac{\sigma_\mu}{\sigma_\delta})$$

which is a probit function that cannot be identified without panel data. Note that P_t' is a measure of the size of the hard-core "permanent poverty" population. It is a cross sectional measure in the following sense: If the poverty line and the set of individual characteristics (and thus b_{it}) remained unchanged over time at their level in the current year, P_t' percent of the population would have expected (permanent) earnings below the poverty line. Of those individuals observed in poverty in any year, P_t'/P_t are expected to remain in poverty on average and the remainder are expected to be out of poverty on average.

The aggregate joint probability of poverty sequences for any two years, t and τ, is given by the mean of the corresponding individual probabilities. For example,

$$(3.13) \qquad P_{t,\tau} = E(\phi_{t,\tau}) = \int_{-\infty}^{+\infty} F(b^*_{it}, b^*_{i\tau}; \gamma^{|t-\tau|}) \cdot \frac{1}{\sigma_\delta\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{\delta}{\sigma_\delta})^2} d\delta$$

Alternatively, the aggregate earnings covariance structure for pairs of years given by equation (2.4) and the relationships (3.11) for each year may be used directly to obtain

(3.14) $$P_{t,\tau} = F(b_{it}, b_{i\tau}; \rho+(1-\rho)\gamma^{|t-\tau|})$$

(3.15) $$P_{\sim t,\tau} = F(-b_{it}, b_{i\tau}; -(\rho+(1-\rho)\gamma^{|t-\tau|}))$$

and so forth.

Similarly, it is straightforward to generalize the aggregate probability statements to a sequence of earnings states for K arbitrary years. Using H_t and J_t defined in equations (3.5) and (3.6), for any K contiguous years

(3.16) $$P_{\underset{\sim}{H}} = F((-1)^{H_1}b_{it_1},\ldots,(-1)^{H_K}b_{it_K}; \frac{1}{\sigma_\mu^2}\underset{\sim}{J}'\underset{\sim}{\Sigma}\underset{\sim}{J})$$

where $\underset{\sim}{\Sigma}$ is defined in equation (2.8). Clearly these mean probabilities could be estimated by a K-variate probit function with the independent variables entering through the $b_{it} = Y^*_{it} - \underset{\sim}{X}_{it}\underset{\sim}{\beta}$ terms. When no serial correlation is present at the individual level ($\gamma = 0$), individual probabilities are independent around an individual mean from period to period (and thus are a product of cumulative normals), but aggregate probabilities are not, even for observationally identical individuals, because of population heterogeneity. This is the special case of population heterogeneity studied by Heckman and Willis [1975] in the context of conception probabilities.

Conditional Probabilities and Relationship to Markov Models

Using the probability statements for single years and pairs of years developed above, it is easy to derive expressions for k-step poverty transition matrices for a given individual (δ specified) or for a group of observationally identical individuals (δ unspecified). The typical element in a k-step transition matrix is the conditional probability of being in (out of) poverty in period $\tau = t + k$ given in (out of) poverty in period t.

Individual conditional probabilities of, say, being in poverty in given poverty (not poverty) in t are of the form

$$\text{(3.17)} \qquad \phi_{i\tau|t} = \frac{\phi_{t,\tau}}{\phi_t} \ , \quad \phi_{i\tau|\sim t} = \frac{\phi_{\sim t,\tau}}{\phi_{\sim t}} \ , \quad \text{etc.}$$

Since the serial correlation model implies that the intertemporal distribution of a given individual's earnings is Markovian (see Feller 1971, pp. 94-96), it is not surprising that the individual k-step transition probabilities in our model share some of the same properties as the k-step transition probabilities in a Markov chain model of an individual's poverty transitions. First, the probability that an individual will be in poverty in period τ is larger if he was in poverty in period t than if he was not in poverty in period t. To see this, note that (1) given the absence of serial correlation $\phi_{t,\tau} = \phi_t\phi_\tau$ and $\phi_{\sim t,\tau} = \phi_{\sim t}\phi_\tau$ while (2) positive serial correlation implies that $\phi_{t,\tau} > \phi_t\phi_\tau$ and $\phi_{\sim t,\tau} < \phi_{\sim t}\phi_\tau$. It follows that $\phi_{\tau|t} > \phi_{\tau|\sim t}$ since

$$\text{(3.18)} \qquad \phi_{\tau|t} = \frac{\phi_{t,\tau}}{\phi_t} > \frac{\phi_t\phi_\tau}{\phi_t} = \phi_\tau$$

and

(3.19) $$\phi_{\tau|\sim t} = \frac{\phi_{\sim\tau,\tau}}{\phi_{\sim t}} < \frac{\phi_{\sim t}\phi_{\tau}}{\phi_{\sim t}} = \phi_{\tau} .$$

Second, it is clear that $\phi_{\tau|t}$ is a monotonically decreasing function of $k = |t-\tau|$ while $\phi_{\tau|\sim t}$ is a monotonically increasing function of k. Specifically, the value of conditioning information deteriorates with time so that

(3.20) $$\phi_{\tau|t} \to F(b^*_\tau) = \phi_\tau \quad \text{and} \quad \phi_{\tau|\sim t} \to F(b^*_\tau) = \phi_\tau \quad \text{as} \quad k \to \infty .$$

Despite these similarities, it should be noted that the transition probabilities in our model are not Markovian because the individual's probability of poverty in period τ is affected not only by his poverty state in period t but also by his entire history of poverty states prior to t. This is an example of the well known fact that a process that is Markovian in a given state space is not generally Markovian when states are aggregated.

In a group of observationally identical individuals, those who are in poverty in t tend to be selected for low values of δ and conversely for those out of poverty in t. Thus, in a heterogeneous population, knowledge of an individual's poverty state at time t provides information about his chances of being in poverty at all subsequent times.

The appropriate k-step transition probabilities for a group of observationally identical individuals are equal to the ratio of the aggregate joint probability of poverty (not poverty) in t and τ divided by the aggregate probability of poverty (not poverty) in t. For example, $P_{\sim\tau|t} = P_{t,\sim\tau}/P_t$ is the probability of being out of poverty in τ given poverty in t.

As a result of serial correlation, $p_{\tau|t}$ monotonically decreases and $p_{\tau|\sim t}$ monotonically increases as $k = |t-\tau|$ increases. However, unlike the case for a given individual with a fixed value of δ, the value of the conditioning information (i.e., poverty status in t) does not erode completely as k approaches infinity. For example,

$$(3.22) \qquad p_{\tau|t} = \frac{F(b_t, b_\tau; \rho + (1-\rho)\gamma^{|t-\tau|})}{F(b_t)} \rightarrow \frac{F(b_t, b_\tau; \rho)}{F(b_t)} > F(b_\tau) = P_\tau$$

$$\text{as } k \rightarrow \infty .$$

Similarly, $P_{\tau|\sim t} < P_\tau$ for all $k = 1,\ldots$. It follows that a random individual chosen from among those who are in poverty in a given year will have a greater chance of being in poverty in all subsequent years than an observationally identical individual chosen randomly from among those not in poverty in that year.

Clearly, these transition matrices can be easily generalized to any number of earnings states using the parameter estimates of the earnings function and its error structure. In addition, the conditional probabilities allow individual characteristics to vary across people and over time and also allow the poverty definition to change.

Finally, joint and conditional probability statements can easily be generalized to K arbitrary years. The probability of any sequence of events (in or out of poverty) over years $t_1, t_2,\ldots,t_K$ is given by an equation analogous to the two year equations with the intersection of the corresponding events expanded to K periods. For example, the probability of being in poverty in year t given in poverty the $K - 1$ previous periods (or similarly for any other arbitrary sequence) is given by

$$(3.24) \qquad P_{t_1|t_2,\ldots,t_K} = \frac{P_{t_1,t_2,\ldots,t_K}}{P_{t_2,\ldots,t_K}}$$

All information about the relative size of permanent, serially correlated and transitory components enters the determination of these probabilities through the correlation matrix $\underset{\sim}{\Sigma}$ while the size of total variation σ^2_μ affects the arguments b_{it}.

4. EMPIRICAL POVERTY DISTRIBUTIONS AND EARNINGS DYNAMICS

In this section we briefly illustrate some of the concepts developed in section 3, using parameter estimates from section 2. Basic patterns are illustrated using black/white comparisons. Because our purpose here is to be illustrative, the probability concepts are explored empirically only for the crudest earnings function in which no explanatory variables except time dummies enter. This form of the earnings function puts the entire analysis in its worst possible light while emphasizing its power to predict the qualitative features of the dynamics of poverty for individuals and groups. This means that many interesting questions related to measured determinants of earnings are not explored at all.

It is important to introduce the caveat that all of the probability statements predicted from the model assume normality of δ and η. While this assumption is not necessary it is quite convenient analytically. Since the probability statements reported in this paper relate to log earnings, it it is appropriate to note that the actual distributions of Y_{it} for both blacks and whites are leptokurtic and slightly negatively skewed relative to normal curves with the same mean and standard deviation.[17] Much of the discrepancy between actual and predicted probabilities can be traced to non-normality.

We begin by comparing actual and predicted poverty sequences for black and white aggregates. These probabilities include single-year poverty probabilities, joint and conditional probabilities for three adjoining years and k-step transition probabilities for up to seven years. We also compare the fraction of blacks and whites in permanent poverty (i.e., whose expected value is the poverty state). The relationships between individual probabilities and aggregate probabilities are then explored. First, we illustrate the distribution of individual single-year and two-year joint poverty probabilities as a function of the permanent component δ and then show the distribution of such probabilities across individuals. Differences in probabilities between an individual chosen at random from a population and a hypothetical representative individual with $\delta=0$ are used to illustrate the effects of population heterogeneity.

Actual and Predicted Poverty Sequences

In our empirical illustrations, we define an individual from the Income Dynamics panel to be in poverty in year t if his earnings, Y_{it}, fall below an arbitrary poverty line, Y_t^*, which is equal to one half median earnings of male workers reported in the corresponding year in the Current Population

Survey. Since the estimated parameters $\hat{\sigma}_\delta^2$, $\hat{\sigma}_\eta^2$ and $\hat{\gamma}$ are all assumed to remain constant over time and no explanatory variables except time dummies are considered, the only source of temporal variation in predicted probabilities in our model is caused by time variation in $\overline{Y}_t - Y_t^*(=\sigma_\mu b_{it})$ where $\overline{Y}_t$ is mean log earnings of whites or blacks in the Income Dynamics panel.[18] Economy-wide productivity change tends to increase both $\overline{Y}_t$ and Y_t^*; however, since the Income Dynamics sample also increases its average level of labor force experience over time, $\overline{Y}_t$ tends to increase somewhat faster than Y_t^*. Accordingly, the poverty threshold, $\overline{Y}_t - Y_t^*$, which is tabulated in Table 2 for blacks and whites, tends to rise from 1967 to 1973.

The actual $(\overline{p}_t)$ and expected (p_t) fractions of blacks and whites in poverty in each year 1967-73 are presented in Table 2, along with the percentage in permanent poverty (p_t').[19] The expected aggregate poverty probabilities are calculated using the probit function in equation (3.12) and the parameter estimates of the log earnings model for whites and blacks in Table 1. According to the predicted probabilities, p_t, a random black is over four times as likely to be in poverty in a given year as is a random white. In large part, this is caused by lower mean earnings for blacks, but to some extent it is also the result of the higher variance of black earnings. For example, in 1970 the black poverty probability is .089. If blacks had the same earnings variance as whites, this probability would have been .066. The decline of the white poverty probability from .028 to .019 and the black probability from .138 to .083 over the period 1967-73 is the result of the growth of mean earnings in the Income Dynamics sample relative to median male

Table 2. Mean Poverty Probability (p_t), Actual Fraction in Poverty ($\overline{p}_t$), Fraction in Permanent Poverty (p_t') and Threshold Values ($Y_t - Y_t^*$) by Year and by Race

YEAR	WHITES				BLACKS			
	$Y_t - Y_t^*$	p_t	$\overline{p}_t$	p_t'	$Y_t - Y_t^*$	p_t	$\overline{p}_t$	p_t'
1967	1.031	.028	.028	.012	.661	.138	.126	.113
1968	1.052	.025	.026	.010	.702	.124	.097	.100
1969	1.062	.024	.026	.010	.734	.113	.097	.090
1970	1.070	.024	.029	.009	.817	.089	.058	.068
1971	1.090	.022	.037	.008	.776	.101	.068	.078
1972	1.101	.021	.029	.008	.840	.083	.078	.062
1973	1.115	.019	.029	.007	.841	.083	.107	.062

earnings in the CPS. Comparing p_t and $\overline{p}_t$ in Table 2, it is apparent that our model predicts the fraction of whites in poverty fairly well, but over-predicts black poverty to a moderate degree.

In Table 3, we present for whites and blacks the joint aggregate probabilities of the eight possible sequences of poverty and nonpoverty over the three-year period 1967-1969 together with the actual proportions following these sequences. The joint probabilities are calculated by evaluating a trivariate normal function according to (3.17) and are reported using the notation $P(H_{1967}, H_{1968}, H_{1969})$ where H_t equals one if not poverty in year t and zero if poverty.[20] Again, as in the case of single-year probabilities, the model is a better predictor for whites than for blacks.

It is easy to see the extent to which population heterogeneity and serial correlation (i.e., $\sigma_\delta^2 > 0$, $\gamma > 0$) increase the probability of persistence in poverty or nonpoverty for a random individual by comparing the values of P(000) and P (111) in Table 3 with the product of single-year poverty (non-poverty) probabilities in Table 2. For whites, the probability of three years of poverty, P(000)=.0076, is over 100 times as large as the product of single-year poverty probabilities, while the probability of three years out of poverty, P(111)=.950, is only modestly larger than the product of probabilities (.925). The corresponding comparisons for blacks are P(000)=.067 compared to .001 and P(111)=.807 compared to .726. Despite this persistence, the three-year probabilities do indicate considerable mobility in and out of poverty. Only 15 percent of whites and 35 percent of blacks expected to be in poverty at any time during 1967-1969 are expected to be in poverty in all three years.

Given heterogeneity and serial correlation, knowledge of an individual's

Table 3. Predicted and Actual Probabilities of Three-Year Sequences of Poverty and Nonpoverty, 1967-1969, by Race

SEQUENCE*	WHITES (N=1238)		BLACKS (N=519)	
	Predicted	Actual**	Predicted	Actual
P(111)	.9495	.9433 (982)	.8074	.8252 (85)
P(110)	.0098	.0125 (13)	.0207	.0291 (3)
P(101)	.0083	.0125 (13)	.0194	.0097 (1)
P(100)	.0044	.0038 (4)	.0142	.0097 (1)
P(011)	.0124	.0144 (15)	.0360	.0388 (4)
P(010)	.0027	.0038 (4)	.0120	.0097 (1)
P(001)	.0053	.0038 (4)	.0237	.0291 (3)
P(000)	.0076	.0058 (6)	.0665	.0485 (5)

*The probabilities are of the form $P(H_{1967}, H_{1968}, H_{1969})$ where $H_t = 1$ if not poverty in t and 0 if poverty. For example, P(101) is the probability of the sequence from 1967 to 1969 of not poverty, poverty, not poverty.

**The number of individuals following each sequence is reported in parentheses.

poverty history is very useful in predicting his poverty status in a given year. This is illustrated by considering predicted and actual aggregate poverty probabilities for whites and blacks in 1969 (a) with no information about poverty history, (b) conditional on poverty status in 1968 and (c) conditional on poverty status in 1967 and 1968. These probabilities are presented in Table 4 where, for example, $P(69|\sim 67, 68)$ is the probability of poverty in 1969 given not poverty in 1967 and poverty in 1968. We shall confine our discussion to predicted probabilities for blacks, since it is apparent from inspection of panels (b) and (c) in Table 4 that similar patterns of conditional probabilities hold for blacks and whites and because predicted and actual conditional probabilities are fairly close for both groups.

A random black in 1969 is predicted to have an 11 percent chance of poverty. If he had been in poverty in 1968, his poverty probability in 1969 is increased to .65, while it would be only .04 if he had been out of poverty in 1968. Having been in poverty in both 1967 and 1968 increases the probability of poverty in 1969 to 74 percent, while being out of poverty in the past two years reduces it to 2.5 percent.

These differences between conditional and unconditional probabilities are the result of the combined effects of heterogeneity and serial correlation. The effect of serial correlation itself can be isolated by comparing $P(69|\sim 67, 68)$ and $P(69|67, \sim 68)$. Aside from slight trend effects, these probabilities would be equal in the absence of serial correlation. With serial correlation, poverty in 1969 is more likely for those who were in poverty more recently. This is borne out by the fact that $P(69|\sim 67, 68) >$ $P(69|67, \sim 68)$ for blacks and whites for both predicted and actual probabilbilities, although for whites the difference in actual probabilities is small.

Table 4. Aggregate Poverty Probabilities in 1969 Conditional on Alternative Poverty Histories by Race

PROBABILITY CONCEPT*	WHITES		BLACKS	
	Predicted	Actual	Predicted	Actual
(a) Unconditional				
P(69)	.024	.026	.113	.097
(b) Conditional on Past Year				
P(69|68)	.469	.370	.652	.600
P(69|∿68)	.013	.017	.037	.043
(c) Conditional on Past Two Years				
P(69|67,68)	.589	.600	.737	.625
P(69|∿67,68)	.346	.235	.423	.500
P(69|67,∿68)	.179	.211	.250	.200
P(69|∿67,∿68)	.010	.013	.025	.034

*∿ indicates not poverty. For example, P(69|67,∿68) is the probability of poverty in 1969 given poverty in 1967 and not poverty in 1968.

The extent to which knowledge of an individual's poverty status in year t continues to provide information about his probability of being in poverty in a subsequent year $\tau = t + k$ may be judged from the actual and predicted k-step transition probabilities presented in Table 5 for t=1967 and τ=1968,...,1973. It is apparent from Table 5 that the risk of poverty to those who were in poverty in 1967 remains considerably larger than the risk to those who were out of poverty in 1967 as much as six years later. This is true for predicted and actual probabilities among both whites and blacks and is chiefly the result of heterogeneity.

Variation in $p_{\tau|t}$ and $p_{\tau|\sim t}$ as τ increases reflects two factors: (1) a decrease in the unconditional poverty probability caused by an increase in mean earnings relative to the poverty line and (2) a decrease in the informational value of serially correlated components in the conditioning year, 1967, as the distance from that year increases. Both factors tend to decrease $p_{\tau|t}$ as τ increases, while the first tends to be offset by the second for $p_{\tau|\sim t}$. Evidently, in the latter case, the two factors are of approximately equal strength since there is little trend in predicted values of $p_{\tau|\sim t}$.

The Distribution of Poverty Probabilities Among Individuals

We now turn to the distribution of individual poverty probabilities that underlie the aggregate probabilities that have been explored to this point. First consider an individual's single-year probability or poverty (using parameter estimates for the total sample from line 1 of Table 1). This probability is clearly a function of the permanent component δ and its

Table 5. Predicted and Actual k-step Poverty Transition Probabilities for Whites and Blacks

	YEAR (τ)					
	1968	1969	1970	1971	1972	1973
A. $P_{\tau\|t}$ = Pr (poverty in τ|poverty in 1967)						
WHITES						
Predicted	.458	.367	.331	.303	.290	.276
Actual	.345	.345	.276	.241	.241	.276
BLACKS						
Predicted	.652	.568	.468	.503	.438	.436
Actual	.615	.462	.385	.231	.385	.462
B. $P_{\tau\|\sim t}$ = Pr (poverty in τ|not poverty in 1967)						
WHITES						
Predicted	.013	.015	.015	.013	.013	.012
Actual	.017	.017	.022	.031	.023	.022
BLACKS						
Predicted	.039	.041	.029	.036	.026	.026
Actual	.022	.044	.011	.044	.033	.056

empirical counterpart is readily calculable by equation (3.2). For example, the probability of being below one half the CPS median earnings ($3334) is one-half for a person with δ 2.2 standard deviations below average ($\delta = -1.04$ and $\sigma_\delta = .47$) and is one-fourth for a person 1.4 standard deviations below average ($\delta = -.67$). The average person ($\delta = 0$) has a .2 percent chance of falling into poverty compared to a 3 percent chance for an individual chosen at random.

The probability of being in poverty in both 1970 and 1971 is similarly illustrated and is also a monotonic function of δ . This probability is given by equation (3.3). This is substantially less than the single-year probabilities but greater than the product of the two single-year probabilities because of serial correlation. For example, consider the person with $\delta = -1.04$ so that $\phi_t(\delta) = .5$. If the years were independent, the joint probability of poverty in both years would be .25. It is instead .31. As years become further apart, the probability approaches .25. The joint probability of poverty in both 1970 and 1973 ($|t-\tau| = 3$) is .26. The individual conditional probability of poverty in 1971 given that this person was in poverty in 1970 is correspondingly .62 rather than .5. This is solely the result of serial correlation, since it is the same person (same δ). The conditional probability of poverty in 1973 given poverty in 1970 is .52.

The joint probability of being in poverty in 1970 but not in poverty in 1971 is not a monotonic function of δ. The joint probability peaks at $\delta = -1.04$. For $\delta > -1.04$, the probability of poverty is small in each year and conversely for $\delta < -1.04$.

The distribution of single year poverty probabilities among individuals is derived directly from the distribution of δ and is stated in equations (3.9) and (3.10). Here we illustrate these distributions for blacks and whites using separate estimates of the components of variance for each group from Table 1.

Consider the distribution function, panel B of Figure 2, and density function, panel A, of poverty probabilities among blacks and whites for 1970. Dispersion in poverty probabilities among individuals within each group is the result of variation in the permanent component, δ. If $\sigma_\delta^2 = 0$, all individuals would have identical poverty probabilities, while if $\sigma_v^2 = 0$, individuals would have either a unitary or zero poverty probability depending on whether δ is greater or less than b_{it}. The calculated distributions presented in Figure 4 lie between these extremes because both transitory and permanent variances are positive.

The differences between aggregate poverty probabilities and probabilities for a representative individual are illustrated by comparing the means and medians of these distributions. The mean of the distribution of poverty probabilities is equal to the proportion expected to be in poverty, which is .024 for whites and .089 for blacks in 1970 (see Table 2). Given normality, the median (i.e., representative) individual has a value of $\delta = 0$. Hence, the median of the poverty probability distribution is $\phi_{it} = F(b^*_{it})$ where $b^*_{it} = 0$.[21] In 1970, both the median white and the median black had a negligible poverty probability of .001. This implies that the representative person is almost completely insulated from poverty.

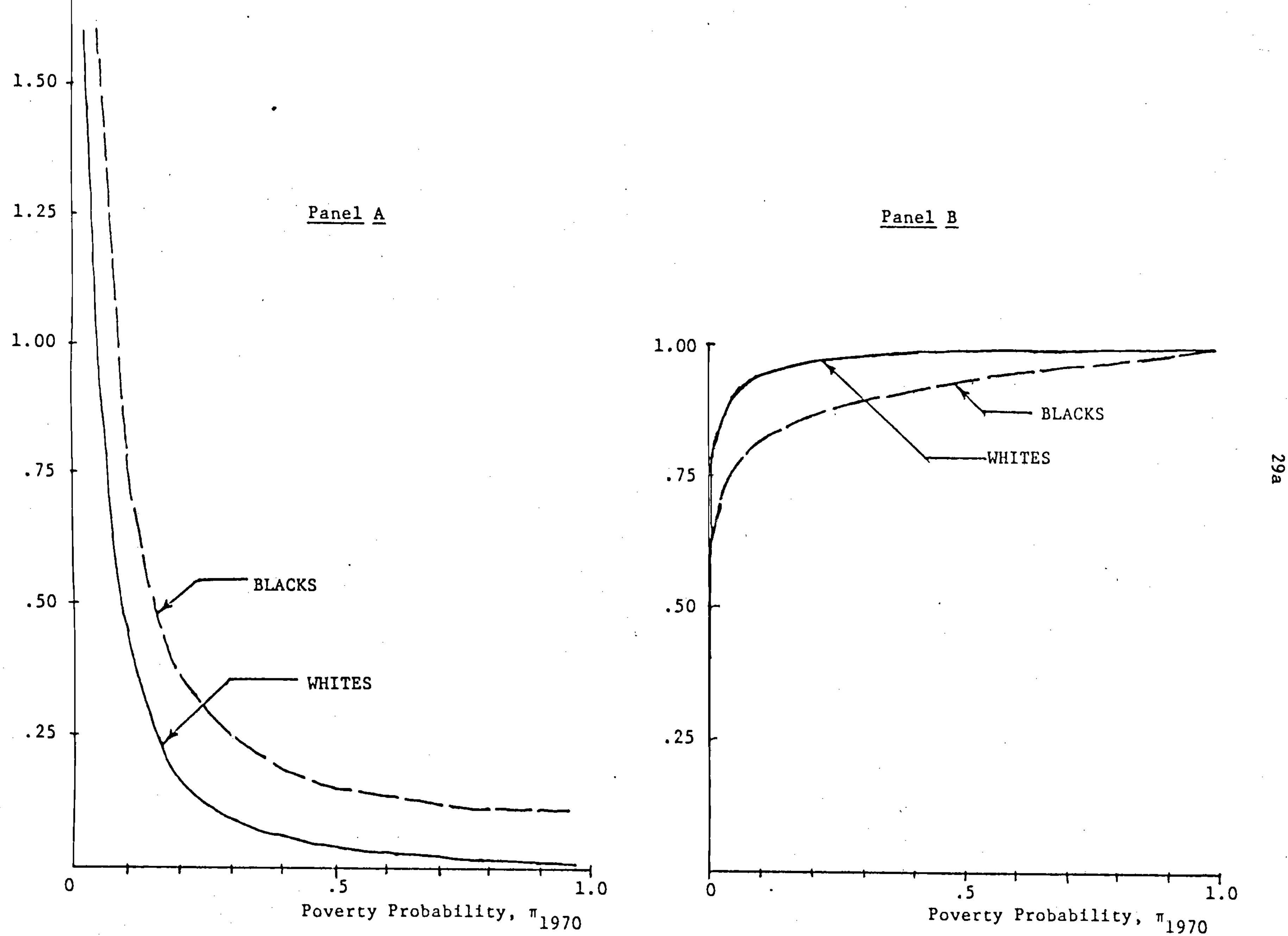

Figure 2. Density Function (Panel A) and Cumulative Density Function (Panel B) of Poverty Probabilities among Individuals in 1970 by Race.

5. SUMMARY AND CONCLUSIONS

In this paper, we present a methodology in which average life cycle earnings growth and the dynamics of the distribution of earnings - viewed either as a continuous distribution or in terms of mobility across a set of discrete earnings classes - can be analyzed within a common econometric framework using longitudinal data. The basic ingredient of our approach is an empirical (log) earnings function with an error structure that allows for permanent differences among individuals due to unmeasured variables and for first order serial correlation in the transitory components of a given individual's time-series of earnings. We call this error structure the "autocorrelated individual component model." Assuming normality of the permanent and transitory components, the intertemporal distribution of log earnings among individuals (holding measured variables constant) is multivariate normal with a correlation structure determined by the share of permanent variance in total variance and the degree of serial correlation. Earnings mobility is then analyzed by deriving the probability statements implied by the earnings function for arbitrary time sequences of earnings states (e.g. whether earnings are above or below an arbitrary poverty line) for a given individual (i.e. holding the permanent component constant) or for a group of individuals. The distribution of poverty probabilities across observationally identical individuals is also derived.

The methodology is illustrated using seven years of data on male earnings from the Michigan Income Dynamics Sample. The autocorrelated individual component model is estimated separately for blacks and whites, as well as for the total sample, using three successively more comprehensive sets of explanatory variables. The simplest model (no explanatory variables except time dummies) indicates that 73.1 percent of total variance in log earnings represents permanent earnings differences. Of the remaining 26.9 percent stochastic variation, 22.5 percentage points are due to purely stochastic variation, and 4.4 to serial correlation.

The variance components and serial correlation coefficients are of roughly similar magnitude for blacks and whites. The permanent component for blacks is about 44 percent larger than for whites, while their transitory component is about the same. Some caution concerning these and other racial comparisons is in order because of the relatively small number of blacks in our sample.

Explanatory variables in the more complex equations tend to leave the size of the permanent and transitory variance components unchanged but reduce the unmeasured permanent variance. For example, schooling, experience, and race explain 33 percent of observed annual earnings variation, but they explain 44 percent of the permanent earnings variation. Schooling, experience, race and the permanent component still explain 73 percent of total earnings variation. Within racial groups, schooling and experience alone explain 51 percent of the permanent component for blacks and 40 percent for whites.

We began this paper with several questions concerning the extent to which poverty is a permanent or transient status and, more broadly, the degree to which the distribution of earnings is characterized more by mobility or stratification. Since the analysis in this paper is confined to males who had earnings in a sequence of years, it cannot deal fully with such questions. A more complete analysis of poverty would consider family income, variations in family composition over time, unemployment and a variety of other issues. However, within their limitations, our model's implications for the mobility of blacks and whites into and out of poverty do suggest some tentative answers to these questions.

The poor are different from the non-poor. Those in poverty in a given year have permanently lower earnings than those not in poverty and are fifteen to twenty-five times as likely to be in poverty as much as six years later. Moreover, these differences are not solely the result of measured characteristics such as race, schooling and experience -- these variables explain only about half of permanent earnings variation with the remainder due to unmeasured factors. Finally, although about 2.5 percent of whites and 9 percent of blacks are predicted to have earnings below the 1970 poverty line, the representative (i.e. median) person of either race had a negligible chance of falling into poverty in that year.

While the poor are different in the sense just described, it would be misleading to conclude that poverty is a permanent status. We find that of those individuals in poverty in a given year, about 55 percent of whites and 35 percent of blacks will be out of poverty in the following year. Another indication of mobility is that only 15 percent of whites and 35 percent of blacks who fall into poverty at some time during the three year period from 1967-70 are expected to be in poverty in all three years. It would also be misleading to conclude that our findings support the concept of a "culture of poverty" which is qualitatively distinct from the social and economic environment in which the majority of persons operate.[22] Rather, our findings simply indicate that the majority of cross-section earnings variation is due to permanent rather than transitory factors. As a result, there is a considerable tendency for individuals to retain their position in the earnings distribution over time whether this position is in the lower, upper or middle portions of the distribution.

FOOTNOTES

[1]Research for this paper was supported in part by grants to the National Bureau of Economic Research by the National Science Foundation (Grant SOC71-03783 A04) and the Hoover Institution. This draft benefited from the helpful comments of A. S. Goldberger at the June 1976 meeting of the Econometric Society. We wish to thank Louis Garrison and Barbara Williams for computational assistance. Responsibility for remaining errors is ours. This research is not an authorized publication of NBER because it has not yet been accorded the full review given to official NBER publications.

[2]See, for example, Rosen [1975] for an explicit statement of this view as a justification for using age-cell means rather than individual data to estimate his earning model. Similar approaches are taken by Heckman [1976], Haley [1976], Ghez and Becker [1975] and others.

[3]See, however, Lillard [1977] and Lillard and Weiss [1977] for examples of the potential value for human capital models of the additional information contained in longitudinal as compared to cross-section or successive cross-section data. Other studies which propose models similar in certain respects to the earnings model we present in Section 2 include Friedman and Kuznets [1954, pp. 352-64], David [1971], Fase [1976], Cooley, McGuire and Prescott [1976] and Hause [1977].

[4]This methodology could be reversed. The reverse methodology has been used in past and current work by James Heckman and Robert Willis on the analysis of panel data on discrete events, such as pregnancy (Heckman and Willis, 1975) and female labor force participation (Heckman, 1977). This reverse methodology could be applied to poverty dynamics as follows. Assuming the error components are normally distributed, sequential observations of poverty and

nonpoverty could be used to estimate the parameters of the earnings function, the level of the poverty line and the covariance structure of unmeasured components (up to a factor of proportionality) using multivariate probit analysis. While this approach has no appeal when earnings and the poverty line can be observed directly, it provides a fruitful framework for the analysis of the timing or sequence of discrete events such as female labor force participation in which the underlying determinants of behavior such as market wages and the shadow price of nonmarket time either cannot be observed or can be observed for only part of the population (e.g. working women).

[5]See Singer and Spilerman (1976) for an excellent survey of Markov and related mobility models.

[6]It clearly reduces to (1) the simple variance component model when $\gamma = 0$, and reduces to (2) a replicated (blockwise) serial correlation model when $\sigma_{\delta}^{2} = 0$.

[7]Alternatively, this error structure may be interpreted as one assuming an infinite history of random shocks.

[8]The SEO sample had been drawn such that all the SEO families chosen by the Survey Research Center for the Income Dynamics Study "...had incomes in 1966 equal to or below twice the federal poverty line at that time. The selection formula was $2,000 + N($1,000) where N is the number of individuals in the family," (Morgan, 1974, p. 2). An earlier version of this paper contained estimates from a sample which included the SEO households. We are grateful to Roger Gordon for calling our attention to the non-random nature of the SEO sample.

[9]Alternative estimates of the earnings function coefficients, β, for the simple function were the same when estimated by OLS and by maximum likelihood jointly with the other parameters of the model.

[10]For the purposes of estimation, the residual error structure is put in the form of a system of linear structural equations, the parameters of which are amenable to estimation by the LISREL III maximum likelihood computer program of Jöreskog and Sorbom (1976). The LISREL model is exposited in Jöreskog (forthcoming) and its application to this structure is explained in detail in Lillard and Weiss (1977).

[11]The regression coefficients and empirical covariance matrices for each set are presented in an appendix available from the authors.

[12]It is probably worth noting at this point that similar magnitudes of these components have been found in other panel data for similar simple earnings functions. For a longitudinal sample of American scientists in the NSF Registry observed over the decade 1960-70, Lillard and Weiss [1976] estimated that $\hat{\rho}$ = .67 and $\hat{\gamma}$ = .70 for gross earnings and that $\hat{\rho}$ = .56 and $\hat{\gamma}$ = .63 for the residuals from a fairly detailed earnings function. The results were stable over a wide range of scientific fields. In another study based on individuals in the NBER-TH sample of W.W. II veterans on whom information is available at intervals from 1943-1969, Lillard (1977) found that 56 percent of earnings variation around an earnings function including age, schooling and ability indices represented permanent differences. This study used a simple variance component model and arithmetic rather than log earnings.

It is also interesting to recall that Friedman's (Friedman, 1967) theory of the consumption function assumes that consumers perform an ex ante decomposition of income variance into permanent and transitory components. Estimates of $\rho = \sigma_\delta^2/\sigma_\mu^2$ (his notation is P_y) obtained by Friedman from cross-section consumption-income data fall within the range of the (ex post estimates of ρ that we obtain from panel data on earnings. For example, his estimates of ρ for urban families is .82 in 1935-36 and .87 in 1941, while corresponding estimates for farm families are .63 and .64 (Friedman, Table 4, p. 67).

[13]Explanatory power is measured here and elsewhere as $1 - \frac{\text{Var(Residual)}}{\text{Var(Total}}$.

[14]This expression was derived by James Heckman for the distribution of women's labor force participation probabilities in a preliminary draft of Heckman and Willis [1977]. The published version of the paper contains the derivation of an analogous expression for $g(\pi)$ for the general case in which the functional forms of the distributions of the permanent and transitory components are not specified.

[15]Note that (3.10) is defined for a group whose measurable characteristics are identical. The aggregate density function over all members of the population is simply the weighted average of these density functions over types of individuals

$$g(\pi_t) = \int_{X_t \varepsilon X} f(\pi_t | X_t)\, f(X_t)\, dX_t$$

The error components themselves may differ with X, e.g., by race.

[16] Or equivalently, noting that $\sigma_v \cdot b^*_{it} \equiv \sigma_\mu \cdot b_{it} - \delta$,

$$P_t = \int_{-\infty}^{+\infty} F(b_{it} \cdot \frac{\sigma_\mu}{\sigma_v} - \frac{\delta}{\sigma_v}) \frac{1}{\sigma_\delta \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{\delta}{\sigma_\delta})^2} d(\delta)$$

[17]These features of the distribution of log earnings have also been reported in other bodies of data in the U.S. and several European countries. See Mincer (1974, p. 113) for discussion and citations to the literature. While a variety of alternative functional forms for cross-section earnings distributions have been proposed (see Singh and Maddala, 1976, for a recent example), we are not aware of any which have a convenient multivariate form.

[18]Another source of differences between predicted and actual aggregate probabilities is the substantial year to year variation in the variance of log earnings.

[19]Since δ cannot be observed, there is no actual counterpart to the estimates of p_t' in Table 2.

[20]These computations were made using a trivariate normal program kindly supplied to us by Ralph Shnelvar. Currently, k-variate normal programs for $k > 3$ are not available to us. Hence we cannot evaluate joint probability statements for more than three (not necessarily adjacent) years.

[21]Note that the median is less than the mean as long as b_{it} is negative and $\sigma_\delta^2 > 0$, i.e., as long as the distribution is not degenerate.

[22]See Levy (1976) for an excellent discussion of the role of the culture of poverty thesis in the debates about poverty policy during the 1960's and further empirical analysis of mobility in the poverty population.

APPENDIX

Table A1. Means, Standard Deviations and Percentages of Variables

VARIABLES	MEAN*	STD. DEV.
EARNINGS:		
1967	8663.2	5099.6
1968	9543.4	5768.7
1969	10361.9	6177.3
1970	10996.3	6673.5
1971	11648.8	7146.3
1972	12657.2	7360.1
1973	13809.2	8142.6
LOG EARNINGS		
1967	9.078	0.550
1968	9.133	0.547
1969	9.166	0.542
1970	9.164	0.546
1971	9.165	0.589
1972	9.227	0.564
1973	9.257	0.545
AGE IN 1967		
EXP IN 1967	23.33	11.00
SCHOOL	11.96	3.37
RACE	.090	
UNION	.314	
UNEMP. RATE	4.91	2.10
DISTANCE	24.00	17.69
CITYSIZE:		
≤500,000	.292	
100,000-499,999	.225	
50,000-99,999	.127	
25,000-49,999	.067	
10,000-24,999	.124	
<10,000	.165	
MARITAL		
MARRIED	.949	
SINGLE	.018	
OTHER	.033	

VARIABLES	MEAN	STD. DEV.
Region		
NE	.220	
NC	.321	
S	.300	
W	.158	
Occup		
PROF	.169	
MGR	.140	
SELF-EMP	.063	
CLERK	.107	
CRAFTS	.224	
OPERATIVE	.165	
LABORER	.067	
FARMER	.040	
MISC	.020	
NOTINLF	.006	
Job Status		
DIFFJOB	.249	
SAMEJOB	.299	
UNREASJOB	.252	
NRJOB	.198	
Employment		
<1 YEAR	.074	
1 YEAR	.098	
2-4 YEARS	.164	
5-9 YEARS	.207	
10-20 YEARS	.223	
>20 YEARS	.099	
NA	.009	
INAPP.OR 0	.126	

*Time varying variables, from union membership on down, are based on 7287 (7N) observations and others are based on 1041 (N) observations.

REFERENCES

Blumen, I., Kogan, M. and P.J. McCarthy, The Industrial Mobility of Labor as a Probability Process, Cornell Studies in Industrial and Labor Relations, Vol. VI, Ithaca, New York, Cornell University, 1955.

Cooley, Thomas F., Timothy W. McGuire and Edward C. Prescott, "A Longitudinal Analysis of the Effects of MDTA Manpower Training Programs." Working Paper #85-75-76. Carnegie Mellon University, July 1976.

David, Martin, "Lifetime Income Variability and Income Profiles," Proceedings of the Social Statistics Section of the American Statistical Association, 1971, pp. 285-292.

Diamond, Peter, Richard Anderson and Yves Balcer, "A Model of Lifetime Earnings Patterns," in Report of the Consultant Panel on Social Security to the Congressional Research Service, Washington: U.S. Government Printing Office, August 1976, Appendix B. pp. 81-119.

Fase, M.M.G., "On the Estimation of Lifetime Income," Journal of the American Statistical Association, (December 1971):686-92.

Feller, William, An Introduction to Probability Theory and its Applications, Volume II (second edition), New York: John Wiley and Sons 1971.

Friedman, Milton and Simon Kuznets, Income from Independent Professional Practice, New York: National Bureau of Economic Research 1954.

Friedman, Milton, A Theory of the Consumption Function, Princeton, National Bureau of Economic Research, 1957.

Ghez, G.R. and G. S. Becker, The Allocation of Time and Goods Over the Life Cycle, National Bureau of Economic Research, 1975.

Gordon, Roger H., "Horizontal Equity and the Federal Income Tax," unpublished Ph.D. dissertation, M.I.T. 1976.

Haley, W.J., "Estimation of the Earnings Profile From Optimal Human Capital Accumulation," Econometrica, Vol. 44, No. 6, November, 1976.

Hart, P.E., "The Dynamics of Earnings, 1963-73," Economic Journal No. 343, September 1976, pp. 551-65.

Hause, John, "The Covariance Structure of Earnings," Annals of Economic and Social Measurement, forthcoming.

Heckman, J.J., "A Life Cycle Model of Earnings, Learnings, and Consumption," Journal of Political Economy, Supplement, August 1976: 511-544.

___________, "Dynamic Models of Female Labor Supply," unpublished paper, University of Chicago, March 1977.

Heckman, J. J. and S. Polachek, "Empirical Evidence on the Functional Form of the Earnings-Schooling Relationship," Journal of the American Statistical Association, June 1974.

Heckman, James J. and Robert J. Willis, "A Beta-Logistic Model for the Analysis of Sequential Labor Force Participation by Married Women," Vol. 85, No. 1, February, 1977.

___________, "Estimation of a Stochastic Model of Reproduction: An Econometric Approach," in Nester E. Terleckyj, ed., Household Production and Consumption, New York, Columbia University Press for NBER, 1975, pp. 99-138.

Joreskog, K.G., "Structural Equation Models in the Social Sciences: Specification, Estimation and Testing," in P.R. Krishnaiah (Ed.) Proceedings of the Symposium on Applications of Statistics. Amsterdam: North Holland Publishing Co. (forthcoming).

Joreskog, K.G. and Sorbom, D., LISREL-III. Estimation of Linear Structural Equation Systems by Maxiumum Likelihood Methods - A Fortran-IV Program, (1976).

Levy, F., "How Big Is the American Underclass?" Berkeley Graduate School of Public Policy Working Paper, revised June, 1976.

Lillard, L. A., "Inequality: Earnings vs. Human Wealth," American Economic Review, March 1977.

Lillard, L. A. and Y. Weiss, "Analysis of Longitudinal Earnings Data: American Scientists 1970-70," NBER Working Paper No. 121, revised January 1977.

McCall, J.J., Income Mobility, Racial Discrimination, and Economic Growth, Lexington Books, 1973.

Mincer, Jacob, Schooling, Experience and Earnings, New York, National Bureau of Economic Research, 1974.

Morgan, J.N., et al., Five Thousand American Families, Ann Arbor, Institute for Social Research, 1974.

Nerlove, M., "A Note on Error Components Models," Econometrica, Vol. 39, No. 2, (March, 1971), pp. 383-396.

Rosen, S., "Toward a Theory of Life Earnings," Journal of Political Economy, Supplement, August 1976: 545-568.

Schiller, Bradley R., "Equality, Opportunity and the 'Good Job'," The Public Interest, No. 43, Spring 1976, pp. 111-120.

Shorrocks, A.F., "Income Mobility and the Markov Assumption," Economic Journal, No. 343, Septamber 1976, pp. 566-78.

Singer, B. and S. Spilerman, "Some Methodological Issues in the Analysis of Longitudinal Survey," Annals of Economic and Social Measurement, Vol. 5, No. 4, Fall 1976, pp. 447-74.

Singh, S.K. and G.S. Maddala, "A Function for the Size Distribution of Incomes," Econometrica, Vol.44, No. 5, September 1976, pp. 963-970.

Weiss, Y. and L.A. Lillard, "Experience, Vintage and Time Effects in the Growth of Earnings: America Scientists, 1960-70," NBER Working Paper No. 138, revised February, 1977.