



ELSEVIER

Journal of Econometrics 96 (2000) 293–356

JOURNAL OF  
Econometrics

www.elsevier.nl/locate/econbase

# An empirical analysis of earnings dynamics among men in the PSID: 1968–1989

John Geweke<sup>a,b,\*</sup>, Michael Keane<sup>b,c</sup>

<sup>a</sup>*Department of Economics, University of Iowa, 121 E. Market Street, Iowa City, IA 52242, USA*

<sup>b</sup>*Department of Economics, University of Minnesota, USA*

<sup>c</sup>*Department of Economics, New York University, USA*

Received 1 December 1997; received in revised form 1 April 1999

---

## Abstract

This study uses data from the Panel Survey of Income Dynamics (PSID) to address a number of questions about life cycle earnings mobility. It develops a dynamic reduced form model of earnings and marital status that is nonstationary over the life cycle. The study reaches several firm conclusions about life cycle earnings mobility. Incorporating non-Gaussian shocks makes it possible to better account for transitions between low and higher earnings states, a heretofore unresolved problem. The non-Gaussian distribution substantially increases estimates of the lifetime return to post-secondary education, and substantially reduces differences in lifetime earnings attributable to race. In a given year, the majority of variance in earnings not accounted for by race, education and age is due to transitory shocks, but over a lifetime the majority is due to unobserved individual heterogeneity. Consequently, low earnings at early ages are strong predictors of low earnings later in life, even conditioning on observed individual characteristics. © 2000 Elsevier Science S.A. All rights reserved.

*Keywords:* Earnings mobility; Panel data; Non-Gaussian disturbances; Markov Chain Monte Carlo

---

## 1. Introduction

This paper models the earnings process of male household heads, using data from the Panel Study of Income Dynamics, 1968–1989. The estimated model

---

\* Corresponding author.

addresses a number of questions about life-cycle earnings mobility. It provides answers to questions such as: What is the probability that a household head with earnings in the bottom quintile of the earnings distribution in one year will still be in the bottom quintile in a subsequent year? What fractions of the variance in lifetime earnings are due to observed heterogeneity, unobserved heterogeneity, and transitory shocks, respectively?

Income mobility has been studied in many previous papers, including McCall (1973), Shorrocks (1976), Lillard and Willis (1978), MaCurdy (1982), Gottschalk (1982), Gottschalk and Moffit (1994). However, we believe that recent advances in econometric methods – in particular, Bayesian inference via Gibbs sampling – make it worthwhile to reexamine this question, because they allow one to estimate much more sophisticated models of the stochastic process for income or earnings than were possible in previous work.

In the classic paper on earnings mobility by Lillard and Willis, the approach is to estimate a standard earnings function, where the dependent variable is log annual earnings and the regressors are education, labor force experience and its square, race, and time effects, and where the error term is assumed to consist of an individual random effect that is normally distributed in the population plus a time-varying normally distributed first-order autoregressive error component. They estimate this model on data from the PSID for male heads of households over the 1967–1973 period. They find that the regressors explain 33% of the variance in log earnings, the random effect accounts for 61% of the error variance, and first-order serial correlation is 0.40.

Some drawbacks of this model are apparent from a comparison of predicted and actual transition probabilities. For instance, the model predicts that, for whites, the probability of being in poverty in 1969 conditional on having been in poverty in 1968 is 46.9%, while the actual sample frequency is only 37%. Thus, the model overstates short-run persistence of the poverty state. Also, the predicted probability of a white person being in poverty in 1969 if he was in poverty in 1968 but not in 1967 is 34.6%, whereas if he was in poverty in 1967 but not in 1968, the predicted probability of being in poverty in 1969 is only 17.9%. The actual sample frequencies of the person being in poverty in 1969 given these past histories are 23.5% and 21.1%, respectively. This again suggests that the model overstates short-run persistence.

A number of possible reasons may explain why the normally distributed random effect plus first-order autoregressive error structure (AR(1)) might overstate short-run persistence and, more generally, fail to fully capture the complexity of observed earnings mobility patterns. One is that the time-varying error term may follow a more complex time-series process than the AR(1) assumed by Lillard and Willis. Another potential problem is that the time-varying error components may not be normally distributed. In fact, Lillard and Willis note that ‘the actual distributions [of log earnings] for both blacks and whites are leptokurtic and slightly

negatively skewed relative to normal curves with the same mean and standard deviation’.

In this paper we focus on the implications of nonnormality of the time-varying error components for estimates of earnings mobility. As described below, it is feasible to undertake Bayesian inference using the Gibbs sampler for models with complex error structures. The latter may have a complex serial correlation structure, with non-Gaussian shocks. In our model the proportion of shock variance due to transitory effects varies with age, for example, and the shape of each of two key shock distributions depends on seven free parameters.

Our work is related to recent work by Horowitz and Markatou (1996), who have developed semiparametric methods for estimating models with random effects plus a transitory error component. They apply this semiparametric approach to a sample of white male workers from the 1986–1987 Current Population Surveys. They find that the transitory component is not normal (it has fatter tails), and show how ‘the assumption that it is normally distributed leads to substantial overestimation of the probability that an individual with low earnings will become a high earner in the future’. In our view, the adoption of a flexible mixture of normals structure for the time-varying errors has some important advantages over a semiparametric approach. In particular, it easily accommodates serial correlation and nonstationarity over the life cycle, and makes fewer demands on the data than do semiparametric methods. The approach of Hirano (1998) to constructing a complex error structure is similar to that taken here. He also uses methods for Bayesian inference much like the ones applied in this paper. However, Hirano uses much smaller, more homogeneous samples than we do, because unlike our model his does not include covariates, and he is unable to use earnings histories that exclude the initial period.

Another reason for reexamining the question of earnings mobility is that much more data is available now than when the classic studies by Lillard and Willis and MaCurdy were done. The PSID now extends over more than 20 yr. Given the objective of distinguishing among alternative serial correlation specifications for the error term, tests based on more than 20 yr of data should have much greater power than ones that use only 7 or 10 yr of data. In particular, one would need a lengthy panel in order to have much hope of distinguishing individual effects from an autoregressive coefficient near one. The model in this paper takes advantage of the longer period, but it also includes data from men who were only observed over very short periods – even as short as one year. In conjunction with a model that permits nonstationarity over the life cycle, the use of all these data required several innovations in methodology, described subsequently.

Finally, we should note that a Bayesian approach has important advantages over classical approaches for studying earnings mobility. Specifically, we can form complete posterior distributions for earnings given any initial state (e.g., parents were black and high school educated) or given any subsequent history

(e.g., respondent obtained a college degree and has a particular earnings history up through age 30). This is, in effect, exactly what Lillard and Willis do, but the posterior distributions they construct are based on classical point estimates. In a Bayesian approach, the posterior distributions are formed by integrating over the posterior distributions of model parameters, thus accounting for parameter uncertainty. In this context, parameter uncertainty is likely to be important, especially since it is difficult to distinguish between individual effects and very strong autoregressive error components. Thus, a prediction of the probability that someone in poverty today will still be in poverty 10 yr from now, based on point estimates of the fraction of variance due to a random effect and the parameters of a complex autoregressive error process, all estimated on only 20 yr of data (not to mention 7 to 10 yr of data), and ignoring the uncertainty in those estimates, does not seem particularly credible.

## **2. The PSID data**

The PSID data set is based on a sample of roughly 5000 households that were interviewed in 1968. Of these, about 3000 were sampled to be representative of the nation as a whole and about 2000 were low-income families that had been interviewed previously as part of the Census Bureau's Survey of Economic Opportunity. The members of these households have been tracked every year since then. People who entered either the original households or split-offs from the original households are also tracked. For example, if after 1968 a child in one of the original households left home to form a new household, then that new household as well as its members are tracked.

The structure of the PSID data is unusual, in that the household is treated as the unit of observation, yet households are unstable over time. Thus, to form a time series of earnings or marital status for an individual in the PSID data, one must determine what household that individual was in during each year of the data (based on unique household identifiers) and then read the individual's earnings and marital status from the relevant household record. For example, if a person was in a particular household in a particular year, and one wants to know the person's earnings, one can determine whether the person was the household head and, if so, read off the earnings-of-household-head variable. Unless the person was the household head in a particular year, data on that individual tend to be scanty.

We use the PSID data for 1968–1989 in our analysis. The full data set contains observations on 38,471 different individuals. We apply several screens to the data. First, we consider only men aged 25–65, and for these men we use only the person-year observations in which the person can be identified clearly as a household head. Our definition of household head is stricter than that in the PSID. Approximately 10% of the males identified as household heads in the

PSID in any given year report they are still students, or that they are keeping house, permanently disabled, in prison, or otherwise institutionalized. We do not count such men as household heads in these periods. Second, we screen out those individuals for whom education or race is unavailable. Third, we drop the observation for the first year a person was a household head, if the earnings information for that year is contained in the data set. We do this because in many cases that is the first year the person works full time, and he may not work the entire year. Such part-year earnings figures may severely understate the person's actual initial earnings potential. Finally, if an individual has missing earnings or marital status observations following his first period of accepted data (due, say, to nonresponse in a particular year), we drop all observations for that person from that point onward. This last screen is convenient, but not essential, because data augmentation methods (see Appendix C<sup>1</sup>) could be used to treat the missing observations as latent variables assuming an independent censoring process. The resulting sample for analysis contains 4766 persons and 48,738 person-year observations. By far the bulk of the sample reduction comes from the first screen: restricting the sample to males aged 25–65 who at some point in the data set are household heads. There are 5267 such individuals in the PSID. The various missing data screens only eliminate 501 of these.

Table 1 reports on the earnings distribution of the analysis sample, conditional on demographics. We define earnings quintiles based on the full sample. In 1967 dollars these are \$3817, \$5786, \$7798 and \$10,454 (to convert to 1998 dollars multiply by 4.88). In Table 1 we report for each of 24 subsamples (two race categories crossed with three education and four age categories) the number of person-year observations in each earnings quintile.

An important aspect of the PSID data is that the earnings questions are retrospective. Most interviews are conducted in March, and the questions refer to earnings in the previous year. Thus, the earnings data in our sample are primarily from 1967 to 1988. We date the observations according to the year of the earnings data, rather than the year of the interview. Consistent with the prior literature on male earnings dynamics and distribution, we drop person-year observations in which reported annual earnings are zero, on the grounds that annual earnings that are truly zero for a male household head are an unusual event. In fact, zero reported annual earnings for males classified as household heads by the PSID are a fairly common event, occurring in approximately 7% of all the person-year observations for males aged 25–65. However, once we apply our stricter definition of a household head, and drop the observations for first time heads, the percent of zeros becomes quite small at most ages. It averages about 1.5% over ages 25–43, increases slowly to about 4% in the late 1940s and

---

<sup>1</sup> Appendices A–G are available through <http://www.econ.umn.edu/~geweke/papers.html>.

Table 1  
Some sample properties of earnings data (full sample)

*Cell counts*

Personal characteristics			Number in earnings quintile					Total
Race	Education	Age	1st	2nd	3rd	4th	5th	
White	< 12	25–34	516	624	461	290	111	2002
White	< 12	35–44	464	562	557	472	276	2331
White	< 12	45–54	582	677	622	536	449	2866
White	< 12	55–65	883	601	454	397	269	2604
White	12–15	25–34	993	1504	1812	1682	891	6882
White	12–15	35–44	457	682	1065	1301	1284	4789
White	12–15	45–54	294	518	703	1063	1052	3630
White	12–15	55–65	374	403	481	580	545	2383
White	≥ 16	25–34	332	319	513	812	843	2819
White	≥ 16	35–44	95	129	230	502	1448	2404
White	≥ 16	45–54	45	90	105	300	1136	1676
White	≥ 16	55–65	83	71	87	211	598	1050
Nonwhite	< 12	25–34	944	524	274	94	23	1859
Nonwhite	< 12	35–44	747	577	420	183	81	2008
Nonwhite	< 12	45–54	881	655	369	201	94	2200
Nonwhite	< 12	55–65	817	405	217	96	59	1594
Nonwhite	12–15	25–34	859	917	716	446	169	3107
Nonwhite	12–15	35–44	156	254	314	283	157	1164
Nonwhite	12–15	45–54	82	111	127	109	72	501
Nonwhite	12–15	55–65	71	39	52	43	47	252
Nonwhite	≥ 16	25–34	49	70	111	84	62	376
Nonwhite	≥ 16	35–44	1	3	34	29	43	110
Nonwhite	≥ 16	45–54	1	9	9	24	18	61
Nonwhite	≥ 16	55–65	21	5	14	9	21	70
Totals			9747	9749	9747	9747	9748	48738

*Sample distributions*

Personal characteristics			Proportion in earnings quintile				
Race	Education	Age	1st	2nd	3rd	4th	5th
White	< 12	25–34	0.258	0.312	0.230	0.145	0.055
White	< 12	35–44	0.199	0.241	0.239	0.202	0.118
White	< 12	45–54	0.203	0.236	0.217	0.187	0.157
White	< 12	55–65	0.339	0.231	0.174	0.152	0.103
White	12–15	25–34	0.144	0.219	0.263	0.244	0.129
White	12–15	35–44	0.095	0.142	0.222	0.272	0.268
White	12–15	45–54	0.081	0.143	0.194	0.293	0.290
White	12–15	55–65	0.157	0.169	0.202	0.243	0.229
White	≥ 16	25–34	0.118	0.113	0.182	0.288	0.299
White	≥ 16	35–44	0.040	0.054	0.096	0.209	0.602
White	≥ 16	45–54	0.027	0.054	0.063	0.179	0.678
White	≥ 16	55–65	0.079	0.068	0.083	0.201	0.570

Table 1 (Continued)

Personal characteristics			Proportion in earnings quintile				
Race	Education	Age	1st	2nd	3rd	4th	5th
Nonwhite	< 12	25–34	0.508	0.282	0.147	0.051	0.012
Nonwhite	< 12	35–44	0.372	0.287	0.209	0.091	0.040
Nonwhite	< 12	45–54	0.400	0.298	0.168	0.091	0.043
Nonwhite	< 12	55–65	0.513	0.254	0.136	0.060	0.037
Nonwhite	12–15	25–34	0.276	0.295	0.230	0.144	0.054
Nonwhite	12–15	35–44	0.134	0.218	0.270	0.243	0.135
Nonwhite	12–15	45–54	0.164	0.222	0.253	0.218	0.144
Nonwhite	12–15	55–65	0.282	0.155	0.206	0.171	0.187
Nonwhite	≥ 16	25–34	0.130	0.186	0.295	0.223	0.165
Nonwhite	≥ 16	35–44	0.009	0.027	0.309	0.264	0.391
Nonwhite	≥ 16	45–54	0.106	0.148	0.148	0.393	0.295
Nonwhite	≥ 16	55–65	0.300	0.071	0.200	0.129	0.300

5% in the late 1950s, and then jumps rapidly to 12.5% at age 62 and to 23% at age 65. The large increases at ages 62 and 65 are due to retirement.

As we discuss in Section 3, in our approach to modeling earnings of male household heads we will adopt the view that the latent earning process runs from age 25 to 65, but that the actual earnings value at any age may be unobserved for various reasons. For example, the latent earnings will be unobserved near the beginning of the life cycle if the person still lives at home or is still a student and is therefore not yet a head. Similarly, we view the earnings process as again becoming unobserved at the end of the life cycle at the point when the person retires. This perspective justifies our ignoring the zero earnings observations for retired men, just as we ignore the zero observations for the young men who are still students or living with parents, and instead viewing latent earnings as unobserved in such cases.

The zeros reported for younger male heads (those under age 62) pose a greater problem. Upon examining the zero observations, we found that 15% of these people were working on the interview date. And, of those not working, only about one quarter reported an unemployment spell of over three months. This makes us suspicious whether many of these people actually had zero labor earnings over the entire previous calendar year. Thus, we suspect that many of these zeros arise due to mis-reporting, and it seems sensible to omit them as well.

Since we treat the stochastic process for earnings as beginning at age 25, we face an initial conditions problem if we do not observe an individual's earnings until an age later than 25. Of the individuals in the sample, only 1728 are observed at age 25, and for these there are 15,604 person-year observations. In part of our analysis, we only use this subsample, which we refer to as the 'young men' sample. This avoids the initial conditions problem. For the full sample, we develop and apply data augmentation methods to the earlier,

missing years. It is worth noting that 569 individuals in the sample have only one year of data, and many others have short records of only a few years of data. Our data augmentation procedure enables us to more than triple the sample size available for inference and to introduce data from later in the life cycle that otherwise could not be used. This procedure can be applied generally in nonstationary models for panel data with partial or interrupted individual records.

### **3. The model**

We model the annual earnings of male household heads between ages 25 and 65. An individual becomes a household head when he ceases to be a dependent; he may be either single or married. For each male in the PSID, our sample begins the year after he became a household head, the year he turns 25, or the year he entered the PSID, whichever is latest. It ends when he left the PSID or turned 65, whichever is later.

In our model the latent process for annual earnings begins at age 25, regardless of the age at which an individual's earnings are first observed. We model earnings at ages greater than 25 as a function of lagged earnings, a set of exogenous personal characteristics (education, age, race, and parents' education), current marital status, individual specific disturbances, and serially correlated shocks. At age 25 annual earnings are a (different) function of the exogenous personal characteristics, and a first-period shock. Realizations of annual earnings from this latent process are observed only when the individual is a household head, is present in the sample, and has been a household head for at least one year. In one variant of the model the first-period and subsequent-period shocks are Gaussian. We refer to this as the 'normal model'. In another variant these shocks are mixtures of three normal distributions and therefore non-Gaussian. We refer to this as the 'mixture model'.

In previous cross-section studies marital status appears to have had a large positive partial correlation with male earnings, even after controlling for human capital variables and other demographic characteristics. Thus, to forecast a man's earnings over all or part of his life cycle it could be important to forecast his marital status as well. We therefore model earnings and marital status jointly.

Marital status is determined in a probit equation. At ages beyond 25 the probit is a function of lagged marital status, lagged earnings, a set of exogenous personal characteristics (education, age, and race), and a serially correlated Gaussian shock. Marital status at age 25 is determined by a probit equation in which the probit is a (different) function of the exogenous personal characteristics and a first-period shock. As with the earnings model, the latent marriage process begins at age 25 regardless of the age at which an individual enters the



data set. Realizations from this process are observed only when the individual is a household head, is present in the sample, and has been a household head for at least one year.<sup>2</sup>

The joint model is fully recursive, with current marital status affecting current earnings, while current earnings do not affect current marital status. This model is applied to a panel of  $n$  individuals,  $i = 1, \dots, n$ . Individual  $i$  is observed in periods  $S_i, \dots, T_i$ , where  $S_i$  and  $T_i$  are determined as just described. Period 1 corresponds to age 25, period 2 to age 26, etc. Because the first-period model is not the same as the model for later periods, and since age appears as a covariate in the later periods, the processes for earnings and marital status are nonstationary. Therefore, if  $S_i > 1$ , the distribution of the first observation on earnings and marital status is an impractically complicated explicit function of the parameters of the model. We avoid this complication by treating the unobserved earnings and marital status in periods  $1, \dots, S_i - 1$  as latent variables, as described in Section 4 and Appendix C. Because of this, it turns out to be harmless to assume that individuals are observed in periods  $1, \dots, T_i$ . With this convention, let  $\Omega_j = \{i: T_i \geq j\}$ , the set of individuals observed in period  $j$ , and let  $N_j$  denote the cardinality of  $\Omega_j$ . The total number of observations is  $N = \sum_{i=1}^n T_i = \sum_{t=1}^T N_t$ .

### 3.1. Earnings model

For  $(t = 1, \dots, T_i; i = 1, \dots, n)$ , further denote

$y_{it}$  = log real earnings of individual  $i$  in period  $t$ ;

$\mathbf{x}_{i1}$  =  $(k_1 \times 1)$  vector of period 1 explanatory variables for individual  $i$   
 $(i = 1, \dots, n)$ ;

$\mathbf{x}_{it}$  =  $(k_2 \times 1)$  vector of period  $t$  explanatory variables for individual  $i$   
 $(t = 1, \dots, T_i; i \in \Omega_2)$ .

---

<sup>2</sup>The marital status data are as of the interview date, while the income data are retrospective. Thus, marital status from March of year  $t$  is paired with income from year  $t - 1$ . It is difficult to pair March of year  $t - 1$  marital status with year  $t - 1$  income information, because a person who was a household head at  $t$  may not have been a head at  $t - 1$ . In this case, time  $t - 1$  information on marital status is often scanty. Note that in either case we must pair point-in-time measures from either March of year  $t - 1$  or March of year  $t$  with annual data that span those dates. Neither approach to dating is 'correct', since both involve an arbitrary pairing of point-in-time with annual measures. Given the data structure of the PSID, it is much more straightforward to pair the March of year  $t$  point-in-time measures with the year  $t - 1$  income data, since both are collected in the same interview.

Table 2  
 Explanatory variables  $\mathbf{x}_{it}$  (earnings model) and  $\mathbf{s}_{it}$  (marriage model)

Variable description	Entry number in			
	$\mathbf{x}_{i1}$	$\mathbf{x}_{it}$	$\mathbf{s}_{i1}$	$\mathbf{s}_{it}$
Indicator (nonwhite)	7	7	1	1
Intercept	9	9	2	2
Education (yr)	10	10	3	3
Age/100	—	11	—	4
Education $\times$ (Age/100)	—	12	—	5
(Age/100) <sup>2</sup>	—	13	—	6
Education $\times$ (Age/100) <sup>2</sup>	—	14	—	7
(Age/100) <sup>3</sup>	—	15	—	—
Education $\times$ (Age/100) <sup>3</sup>	—	16	—	—
Indicator (Married)	8	8	—	—
Indicator (Lagged married)	—	—	—	8
Lagged log earnings	—	—	—	9
Indicator (Father education missing)	1	1	—	—
Indicator (Father education 12 +)	2	2	—	—
Indicator (Father education 16 +)	3	3	—	—
Indicator (Mother education missing)	4	4	—	—
Indicator (Mother education 12 +)	5	5	—	—
Indicator (Mother education 16 +)	6	6	—	—

The model of individual earnings is

$$y_{i1} = \tilde{\beta}' \mathbf{x}_{i1} + \varepsilon_{i1},$$

$$y_{it} = \gamma y_{i,t-1} + (1 - \gamma) \beta' \mathbf{x}_{it} + (1 - \gamma) \tau_i + (1 - \gamma) \phi \varepsilon_{i1} + \varepsilon_{it} \quad (t = 2, \dots, T), \quad (1)$$

$$\varepsilon_{i2} = \tilde{\rho} \varepsilon_{i1} + \eta_{i2}, \quad (2)$$

$$\varepsilon_{it} = \rho \varepsilon_{i,t-1} + \eta_{it} \quad (t = 3, \dots, T),$$

$$\tau_i \stackrel{\text{iid}}{\sim} N(0, \sigma_\tau^2).$$

The vector  $\mathbf{x}_{it}$  used in this study is described in Table 2;  $k_1 = 10$  and  $k_2 = 16$ . The vector  $\mathbf{x}_{i1}$  contains indicator variables for parents' education and the individual's race, and the individual's years of education and current marital status. The vector  $\mathbf{x}_{it}$  ( $t \geq 2$ ) contains these variables and, in addition, a polynomial in education and age, through the first power in education and the third power in age. These are all standard covariates in earnings equations. The functional form of (1) is chosen so that  $\beta_j$  is, to a good approximation, the marginal effect of the corresponding covariate on the unconditional expectation of log real earnings. Thus the polynomial in age and education provides

a model of the effects of aging and education on expected log real earnings that is quite flexible. The first period is taken to be fundamentally different from the remaining periods. Covariates will not have the same effects then as later. Given the dynamic structure, it would be inappropriate to assume that these effects are the same.

The shocks  $\varepsilon_{i1}$ ,  $\eta_{i2}$ ,  $\eta_{it}$  ( $t \geq 3$ ) are mutually independent across both time and individuals. The shocks  $\varepsilon_{i1}$  are identically distributed, as are the shocks  $\eta_{it}$  ( $t \geq 2$ ), but the two do not necessarily have the same distribution. Individual heterogeneity consists of two components. The first-period shock is that portion of first-period earnings that is unanticipated across individuals, conditional on first-period covariates. Part of this shock may be a transitory first-period effect (2), but part of it can also be permanent (the coefficient  $\phi$  in (1)). The mean level of earnings in the dynamic equation (1) is also heterogeneous, by virtue of the shock  $\tau_i$ . The variance of the disturbance vector  $\varepsilon'_i = (\varepsilon_{i1}, \dots, \varepsilon_{iT})$  is a function of the six terms  $\text{var}(\varepsilon_{i1})$ ,  $\text{var}(\eta_{it})$ ,  $\sigma_\tau^2$ ,  $\rho$ ,  $\tilde{\rho}$ , and  $\phi$ : in general a variance matrix for disturbances from any three years corresponds to six values of these parameters, and the fraction of variance due to unobserved heterogeneity (in  $\tau_i$  and  $\varepsilon_{i1}$ ) can range from zero to one and can change smoothly from year to year.

Serial correlation in the distribution of earnings, conditional on covariates, enters the model by means of the parameters  $\gamma$  and  $\rho$ . All of the covariates in  $\mathbf{x}_{it}$ , except current marital status, are either time invariant or are deterministic polynomial functions of time. Thus the only nondeterministic covariate, marital status, provides most of the distinction between  $\gamma$  and  $\rho$ : effects of lagged marital status are present if and only if  $\gamma \neq 0$ . If  $\rho \neq 0$  but  $\gamma = 0$  there is serial correlation in earnings but no lagged impact of marital status on earnings.

In the mixture model the distributions of  $\varepsilon_{i1}$  and  $\eta_{it}$  are each mixtures of three normal distributions; e.g.,

$$\eta_{it} \sim \text{N}(\alpha_{2(j)}, h_{2(j)}^{-1}) \text{ with probability } p_{2j},$$

where  $\alpha_{2(1)} < \alpha_{2(2)} = 0 < \alpha_{2(3)}$ ;  $0 < h_{2(j)} < \infty$  ( $j = 1, 2, 3$ );  $p_{2j} \geq 0$  ( $j = 1, 2, 3$ ); and  $p_{21} + p_{22} + p_{23} = 1$ . (Similarly,  $\varepsilon_{i1} \sim \text{N}(\alpha_{1(j)}, h_{1(j)}^{-1})$  with probability  $p_{1j}$ ). The shock distribution thus belongs to a seven-parameter family in each case. This feature of our model is unusual but important. It turns out that shocks are indeed non-Gaussian, and the mixture of three normal distributions goes far to resolve the puzzle about predicted and actual transitions noted in the introduction. The normal model is a special case of this model, which imposes the constraint  $p_{j2} = 1, p_{j1} = p_{j3} = 0$  ( $j = 1, 2$ ).

The choice of three components for the mixture of normals was based on inspection of the posterior distribution of the  $\varepsilon_{it}$  when the shocks are assumed to be normal. We found through experimentation that a three-component mixture model could approximate the posterior distribution of the  $\varepsilon_{it}$  reasonably well. Given the lack of experience with mixture models this complex, we did not wish to use more components. Based on the results reported below, it seems clear that

experimentation with more components for the mixture of normals is warranted. An alternative is to allow an infinite number of components, and use a Dirichlet process prior for the number of components actually observed in the sample. The latter approach was taken by Hirano (1998), but his model excludes covariates, and allows either random effects or autocorrelation in transitory shocks but not both.<sup>3</sup>

The earnings model has 45 free parameters. It is completed with a prior distribution for these parameters. We choose a prior distribution in the light of two criteria. First, the functional form of the prior distribution should be one that is flexible but also convenient in obtaining the posterior distribution. This relationship between the functional form of the prior and posterior is treated in Appendices 1 and 2 of this paper, and in detail in Appendices A and B. Second, the prior distribution should center about values that are plausible in the context of the earnings and income mobility literature, but should also be diffuse enough to permit all reasonable (and in the process, many unreasonable) departures from these values. A detailed presentation of the prior distribution is made in Appendix 3.1.

One feature of the prior distribution is worth emphasis, for it copes with the interpretation of the effects of age and education on earnings in a way that is also useful in the subsequent presentation of results. The prior distribution for the coefficients of the age-education polynomial is developed by considering the difference between expected log earnings at age  $a_2$  and education  $e_2$ , and expected log earnings at age  $a_1$  and education  $e_1$ , denoted  $G(a_1, a_2; e_1, e_2)$ . Independent, normal prior distributions for  $G(25, 35; 12, 12)$ ,  $G(35, 45; 12, 12)$ ,  $G(45, 55; 12, 12)$ ,  $G(25, 25; 12, 16)$ ,  $G(35, 35; 12, 16)$ ,  $G(45, 45; 12, 16)$  and  $G(55, 55; 12, 16)$  were constructed. Combined with another independent prior distribution for expected log earnings at age 25 and education level 12, these eight distributions imply a joint normal distribution on the coefficients in the polynomial in education (powers 0 and 1) and age (powers 0 through 3). Since individual coefficients in this polynomial have no interesting interpretation, we make use of this convention as well in subsequently reporting posterior means.

Assessment of the sensitivity of the posterior distribution to changes in the prior is useful in interpreting the results. Corresponding to each posterior mean, we report the posterior standard deviation, and the prior mean and standard deviation. This facilitates a quick approximation of the sensitivity of the posterior mean to the prior mean using the familiar relation that is exact in the Gaussian case:  $\bar{\theta} = (h\underline{\theta} + h\theta)/(h + h)$ , where  $\underline{\theta}$ ,  $\theta$ , and  $\bar{\theta}$  are the prior, data, and

---

<sup>3</sup> For further discussions of finite normal mixture models see West and Harrison (1989), Section 12.3.4, and Roeder and Wasserman (1997). On infinite normal mixture models with Dirichlet process priors see Escobar and West (1995) and MacEachern and Mueller (1995).

posterior means, respectively; and  $\underline{h}$  and  $\bar{h}$  are the prior and data precisions. Since the posterior precision is  $\bar{h} = \underline{h} + h$ , the sensitivity of the posterior mean to the prior mean is  $\partial\bar{\theta}/\partial\theta = \underline{h}/(\underline{h} + h)$ , which is simply the ratio of the posterior variance to the prior variance. This expression is invariant to rescaling of  $\theta$ , bounded between 0 and 1, and naturally interpreted as a ratio.

### 3.2. Marital status model

We adopt a dynamic probit specification for marital status. Denote

- $d_{it} = 1$  if individual  $i$  is married in period  $t$   
and  $d_{it} = 0$  if not ( $t = 1, \dots, T_i; i = 1, \dots, n$ );
  - $s_{i1}$  = ( $p_1 \times 1$ ) vector of period 1 explanatory variables  
for individual  $i$  ( $i = 1, \dots, n$ );
  - $s_{it}$  = ( $p_2 \times 1$ ) vector of period  $t$  explanatory variables  
for individual  $i$  ( $t = 2, \dots, T_i; i \in \Omega_2$ );
  - $m_{it}^*$  = Probit (latent) that determines  $d_{it}$  ( $t = 1, \dots, T_i; i = 1, \dots, n$ ).
- The model for marital status is

$$m_{i1}^* = \tilde{\theta}'s_{i1} + \xi_{i1},$$

$$\xi_{i1} \stackrel{\text{iid}}{\sim} N[0, (1 - \lambda^2)^{-1}],$$

$$m_{it}^* = \theta's_{it} + \xi_{it} \quad (t = 2, \dots, T),$$

$$\xi_{it} = \lambda\xi_{i,t-1} + \psi_{it} \quad (t = 2, \dots, T),$$

$$\psi_{it} \stackrel{\text{iid}}{\sim} N(0, 1) \quad (t = 2, \dots, T),$$

$$d_{it} = \begin{cases} 1 & \text{if } m_{it}^* \geq 0, \\ 0 & \text{if } m_{it}^* < 0. \end{cases}$$

The vector  $s_{it}$  used in this study is described in Table 2;  $p_1 = 3$  and  $p_2 = 9$ . The vector  $s_{i1}$  contains an intercept, the individual's education, and a race indicator. The vector  $s_{it}$  ( $t \geq 2$ ) contains these variables and, in addition, lagged marital status  $d_{i,t-1}$  and log real earnings  $y_{i,t-1}$ , and a polynomial in education and age, through the first power in education and the second power in age. As in the earnings model, the specification of the first-period equation is different from the other periods. The most important factor dictating a different structure is that we do not have available lagged earnings for the first period, as explained above. We retain an explicit latent-variable formulation for the model for two reasons. First, this representation is readily amenable to the computational methods outlined subsequently. Second, in extensions and elaborations of this work, we intend to allow for the possibility that shocks to continuous and

discrete variables may be dependent. This possibility is facilitated by the latent-variable representation.

The marital status model has 13 free parameters. It is completed with a prior distribution for these parameters, designed according to the same criteria used in developing the earnings model prior. A detailed presentation of the marital status model prior distribution is made in Appendix 3.2. As in the earnings model it is necessary to cope with the interpretation of the effects of age and education – here, on the marital status probit. The prior distribution for the coefficients of the age-education polynomial is developed by considering the difference between the expected marital status probit at age  $a_2$  and education  $e_2$ , and the expected marital status probit at age  $a_1$  and education  $e_1$ , denoted  $\Delta(a_1, a_2; e_1, e_2)$ . Independent, normal prior distributions for  $\Delta(25, 40; 12, 12)$ ,  $\Delta(40, 55; 12, 12)$ ,  $\Delta(25, 25; 12, 16)$ ,  $\Delta(40, 40; 12, 16)$  and  $\Delta(55, 55; 12, 16)$  were constructed. Combined with another independent prior distribution for the expected marital status probit at age 25 and education 12, these six distributions imply a joint normal distribution on the coefficients in the polynomial in education (powers 0 and 1) and age (powers 0 to 2). Since individual coefficients in this polynomial have no interesting interpretation, we make use of this convention as well in subsequently reporting posterior means.

#### 4. Bayesian inference

This section provides an overview of the methodology for conducting Bayesian inference in the earnings-marital status model. This description assumes familiarity with Bayesian inference and with the Gibbs sampling algorithm for drawing values from a posterior distribution. Accessible introductions to both topics for economists include Chib and Greenberg (1995,1996) and Geweke (1996,1999). More general references are Gelman et al. (1995) and Gilks et al. (1996).

The objective here is to provide an overview of the methods that are described in complete detail in Appendices A, B, and C. To that end, some additional notation is useful. Let  $\mathbf{z}_i$  denote the vector of time invariant or deterministic characteristics of individual  $i$ : i.e., all variables except earnings and marital status. Let  $L_{it}$  be an integer latent variable indicating from which of the three normal distributions the shock  $\varepsilon_{i1}$  (if  $t = 1$ ) or  $\eta_{it}$  (if  $t \geq 2$ ) was drawn. Let  $\mathbf{Y}'_{it} = (y_{i1}, \dots, y_{it})$ ,  $\mathbf{D}'_{it} = (d_{i1}, \dots, d_{it})$  and  $\mathbf{M}^*_{it} = (m^*_{i1}, \dots, m^*_{it})$ . Finally, let  $\theta_E$  denote the  $45 \times 1$  vector of parameters in the earnings model, and  $\theta_M$  the  $13 \times 1$  vector of parameters in the marital status model.

The earnings model outlined in Section 3.1 and described in complete detail in Appendix 1 provides the probability density functions

$$p_E(y_{it} | \mathbf{Y}_{i,t-1}, \mathbf{D}_{it}, \mathbf{z}_i, \tau_i, L_{it}, \theta_E), \quad p_E(\tau_i | \theta_E), \quad p_E(L_{it} | \theta_E).$$

The marital status model outlined in Section 3.2 and described in complete detail in Appendix 2 provides the probability density function and probability function

$$p_M(m_{it}^* | Y_{i,t-1}, M_{i,t-1}^*, z_i, \theta_M), \quad p_M(d_{it} | m_{it}^*).$$

The corresponding prior distributions for each model provide, respectively,  $p_E(\theta_E)$  and  $p_M(\theta_M)$ .

By the standard definition of conditional probability,

$$\begin{aligned} & p\{\theta_M, \theta_E, [\tau_i, (y_{it}, d_{it})_{i=1}^{S_i-1}, (L_{it}, m_{it}^*)_{i=1}^{T_i}]_{i=1}^n \mid [z_i, (y_{it}, d_{it})_{t=S_i}^{T_i}]_{i=1}^n\} \\ & \propto p_E(\theta_E) \prod_{i=1}^n \left\{ p_E(\tau_i \mid \theta_E) \prod_{t=1}^{T_i} [p_E(L_{it} \mid \theta_E) p_E(y_{it} \mid Y_{i,t-1}, \mathbf{D}_{it}, z_i, \tau_i, L_{it}, \theta_E)] \right\} \\ & \times p_M(\theta_M) \prod_{i=1}^n \left[ \prod_{t=1}^{T_i} p_M(m_{it}^* \mid Y_{i,t-1}, \mathbf{D}_{i,t-1}, M_{i,t-1}^*, z_i, \theta_E) p_M(d_{it} \mid m_{it}^*) \right]. \end{aligned}$$

We use a Gibbs sampling algorithm to make draws from this conditional distribution. (More precisely, a Gibbs sampling algorithm is used to construct a Markov chain whose unique invariant distribution is this distribution.) The algorithm proceeds in three groups of steps, detailed in Appendices A, B, and C, respectively.

In the first group of steps, the parameter vector  $\theta_E$  is divided into eight blocks. A drawing is made from each block, conditional on all other parameters and latent variables. Then the individual effects  $\tau_i$  ( $i = 1, \dots, n$ ) are drawn individually and in succession, exploiting their conditional independence. Finally the  $L_{it}$  ( $t = 1, \dots, T_i; i = 1, \dots, n$ ) are drawn in succession, again taking advantage of conditional independence. This completes a set of drawings from the conditional distributions for all parameters and latent variables in the earnings model, given  $(Y_{i,T_i}, \mathbf{D}_{i,T_i})_{i=1}^n$ . The algorithm is described in Appendix A. Details for the parameters of the mixture distribution are given in Appendix F.

In the second group of steps, the parameter vector  $\theta_M$  is divided into two blocks. A drawing is made from each block, conditional on all other parameters and latent variables. Then the probits  $m_{it}^*$  ( $t = 1, \dots, T_i; i = 1, \dots, n$ ) are drawn individually; these are conditionally independent across individuals but not across time periods. This completes a set of drawings from the conditional distributions for all parameters and latent variables in the marital status model, given  $(Y_{i,T_i}, \mathbf{D}_{i,T_i})_{i=1}^n$ .

In the third group of steps, first the unobserved earnings  $(Y_{i,S_i-1})_{i=1}^n$  are drawn. These are conditionally independent across individuals and jointly normally distributed. Then, the unobserved probits and marital statuses  $(\mathbf{D}_{i,S_i-1}, M_{i,S_i-1}^*)_{i=1}^n$  are drawn. These are conditionally independent across individuals, but not across time periods, and so are drawn in succession for each individual. For the sample of young men, all  $S_i = 1$  and this third group of steps is skipped.

Because the shocks  $\varepsilon_{it}$  and  $\eta_{it}$  have mixture of normals distributions, the likelihood function is unbounded. The essential problem is that coefficients on the covariates can be chosen to make the conditional means of several log earnings observations identical to their observed values, given in one of the three normal distributions. As the variance term for that distribution approaches zero, the likelihood function is unbounded. This property of the mixture of normals likelihood function was first demonstrated by Kiefer and Wolfowitz (1956); for an extended discussion, see Titterton et al. (1985, Section 4.3). As a consequence, numerical problems can arise in maximum likelihood algorithms and there is no assurance that any given bounded local maximum of the likelihood function is consistent (Redner and Walker, 1984). Appendix F shows that given conventional gamma priors for the precision terms  $h_{k(j)}$ , the posterior distribution exists, as do moments of bounded functions of interest. Appendix F also shows that the posterior moment of an unbounded function of interest exists if the corresponding prior moment exists after reducing the degrees of freedom parameters in the chi square priors for the precision terms  $h_{i(j)}$  by  $2 + \varepsilon$  ( $\varepsilon > 0$ ). With the exception of four moments noted in Table 3, the first and second moments of all unbounded functions of interest reported here satisfy the latter condition.

The Gibbs sampling algorithm simulates a Markov chain in high-dimensional space. By following all of the steps of the algorithm detailed in Appendices A, B, and C, it can be verified that the probability that this Markov chain will move from any point in this parameter space to any region of the space with strictly positive posterior probability, in exactly one complete step of the algorithm, is nonzero. The chain is therefore ergodic (Tierney, 1994; Geweke, 1996): i.e., if  $E\{g(\theta_E, \theta_M) | [z_{it}, (y_{it}, d_{it})_{t=S_i}^T]_{i=1}^n\}$  exists, then the corresponding sample average of  $g(\theta_E, \theta_M)$  from the posterior simulator converges almost surely to this posterior moment.

Operationally, the Gibbs sampling algorithm produces a file with one record for each iteration. Each record has 58 entries, the parameter values for that iteration. Some posterior moments can be approximated directly from this file by corresponding sample averages of explicit functions of parameters. (One example is the serial correlation parameter  $\rho$  in the earnings model. Another is the difference in unconditional expected log real earnings at ages 35 and 25, given 16 yr of education.) Most of the questions we investigate, however, have to do with properties of the earnings process. To facilitate this investigation, we construct a second file of simulated earnings and marital statuses, based on the Gibbs sampling output file and the personal characteristics of the individuals in the sample. Corresponding to the personal characteristics of each individual in the sample, we randomly select ten sets of parameter values from the Gibbs sampling output file. Then we simulate the model from period 1 (age 25) through period 41 (age 65) and record the simulated path of earnings and marital status in each case. (For details of the simulation procedure, see Appendix D.) The simulated values are then used to approximate the probabilities of various



Table 3

Prior and posterior means and standard deviations for parameters and functions of interest, earnings and marital status models

	Young men		Full sample	
	Prior	Mixed model	Normal model	Mixed model
			Normal model	Normal model
<i>Earnings period 1 covariates</i>				
Father ed missing	0.050 (0.200)	- 0.163 (0.056)	- 0.237 (0.080)	- 0.124 (0.067)
Father ed 12 +	0.050 (0.100)	0.014 (0.023)	0.025 (0.036)	- 0.011 (0.027)
Father ed 16 +	0.100 (0.100)	0.001 (0.036)	- 0.019 (0.051)	- 0.042 (0.041)
Mother ed missing	0.050 (0.100)	0.059 (0.065)	0.081 (0.096)	0.133 (0.105)
Mother ed 12 +	0.050 (0.200)	- 0.021 (0.021)	- 0.006 (0.032)	- 0.030 (0.028)
Mother ed 16 +	0.100 (0.100)	- 0.032 (0.046)	- 0.029 (0.061)	0.023 (0.052)
Nonwhite indicator	- 0.100 (0.100)	- 0.191 (0.024)	- 0.261 (0.035)	- 0.195 (0.030)
Marital status current	0.200 (0.200)	0.072 (0.022)	0.083 (0.029)	- 0.006 (0.048)
Intercept	7.22 (4.00)	7.85 (0.079)	7.86 (0.116)	8.09 (0.158)
Education	0.065 (0.075)	0.050 (0.006)	0.045 (0.009)	0.032 (0.011)
<i>Earnings period t covariates</i>				
Father ed missing	0.050 (0.200)	- 0.196 (0.050)	- 0.139 (0.065)	- 0.040 (0.100)
Father ed 12 +	0.050 (0.100)	- 0.011 (0.021)	0.045 (0.028)	0.085 (0.042)
Father ed 16 +	0.100 (0.100)	0.016 (0.033)	0.010 (0.042)	- 0.060 (0.076)
Mother ed missing	0.050 (0.100)	0.093 (0.067)	0.122 (0.093)	- 0.103 (0.066)
Mother ed 12 +	0.050 (0.200)	- 0.006 (0.018)	- 0.009 (0.025)	- 0.019 (0.040)
Mother ed 16 +	0.100 (0.100)	- 0.086 (0.042)	- 0.046 (0.050)	0.005 (0.089)
Nonwhite indicator	- 0.100 (0.100)	- 0.208 (0.021)	- 0.268 (0.026)	- 0.164 (0.051)
Marital status current	0.200 (0.200)	0.036 (0.008)	0.081 (0.013)	0.050 (0.009)
Earnings age 25, Ed 12	8.00 (4.00)	8.54 (0.024)	8.41 (0.030)	8.51 (0.037)
Earnings age 35 vs. 25, Ed = 12	0.150 (0.100)	0.236 (0.117)	0.245 (0.024)	0.231 (0.029)
Earnings age 45 vs. 35, Ed = 12	0.100 (0.100)	0.109 (0.023)	0.125 (0.034)	0.081 (0.008)
Earnings age 55 vs. 45, Ed = 12	0.050 (0.100)	0.157 (0.116)	0.099 (0.095)	- 0.043 (0.008)
Earnings ed 16 vs. 12, Age = 25	0.260 (0.150)	0.195 (0.026)	0.173 (0.038)	0.294 (0.034)
Earnings ed 16 vs. 12, Age = 35	0.340 (0.200)	0.341 (0.021)	0.450 (0.029)	0.469 (0.038)
Earnings ed 16 vs. 12, Age = 45	0.370 (0.225)	0.374 (0.033)	0.389 (0.055)	0.483 (0.040)
Earnings ed 16 vs. 12, Age = 55	0.400 (0.250)	0.284 (0.197)	0.400 (0.230)	0.447 (0.037)

Table 3 (Continued)

	Young men			Full sample		
	Prior	Mixed model	Normal model	Mixed model	Normal model	Normal model
<i>Properties of first period shock<sup>a</sup></i>						
Mean 1	-3.00 (1.00)	-1.966 (0.233)	—	-2.625 (0.271)	—	—
Mean 2	0.00 (0.00)	0.00 (0.00)	—	0.00 (0.00)	—	—
Mean 3	0.100 (0.100)	0.147 (0.029)	—	0.199 (0.004)	—	—
Standard deviation 1	1.425 (0.102)	1.345 (0.074)	—	1.525 (0.114)	—	—
Standard deviation 2	1.127 (0.081)	0.791 (0.035)	0.774 (0.013)	0.776 (0.033)	—	0.809 (0.016)
Standard deviation 3	0.319 (0.022)	0.329 (0.013)	—	0.321 (0.013)	—	—
Probability 1	0.020 (0.014)	0.058 (0.011)	—	0.046 (0.007)	—	—
Probability 2	0.480 (0.050)	0.337 (0.033)	—	0.385 (0.029)	—	—
Probability 3	0.500 (0.050)	0.606 (0.031)	—	0.568 (0.028)	—	—
P{ < log(0.2) }	0.052 (0.012)	0.041 (0.004)	0.019 (0.002)	0.041 (0.003)	—	0.023 (0.002)
P( < log(0.5) ]	0.149 (0.016)	0.111 (0.005)	0.185 (0.004)	0.113 (0.004)	—	0.195 (0.005)
P[ < log(0.8) ]	0.296 (0.023)	0.250 (0.006)	0.387 (0.002)	0.242 (0.006)	—	0.391 (0.002)
P( < log(0.9) ]	0.367 (0.029)	0.321 (0.009)	0.446 (0.001)	0.309 (0.008)	—	0.448 (0.001)
P( < 0 ]	0.442 (0.035)	0.401 (0.010)	0.500 (0.000)	0.383 (0.009)	—	0.500 (0.000)
P[ > log(1.111) ]	0.476 (0.038)	0.508 (0.012)	0.446 (0.001)	0.529 (0.010)	—	0.448 (0.001)
P[ > log(1.25) ]	0.385 (0.039)	0.402 (0.012)	0.387 (0.002)	0.424 (0.010)	—	0.391 (0.002)
P[[ > log(2) ]	0.149 (0.020)	0.103 (0.006)	0.185 (0.004)	0.110 (0.006)	—	0.195 (0.005)
P[ > log(5) ]	0.038 (0.009)	0.008 (0.001)	0.019 (0.002)	0.008 (0.001)	—	0.023 (0.002)
<i>Properties of tth period shock<sup>a</sup></i>						
Mean 1	-3.00 (1.00)	-0.955 (0.088)	—	-0.899 (0.043)	—	—
Mean 2	0.00 (0.00)	0.00 (0.00)	—	0.00 (0.00)	—	—
Mean 3	0.100 (0.100)	0.064 (0.046)	—	0.066 (0.014)	—	—
Standard deviation 1	1.425 (0.102)	1.313 (0.058)	—	1.284 (0.029)	—	—
Standard deviation 2	1.127 (0.081)	0.574 (0.015)	0.466 (0.003)	0.462 (0.008)	—	0.448 (0.002)
Standard deviation 3	0.319 (0.022)	0.146 (0.003)	—	0.117 (0.001)	—	—
Probability 1	0.020 (0.014)	0.044 (0.006)	—	0.051 (0.003)	—	—
Probability 2	0.480 (0.050)	0.272 (0.009)	—	0.315 (0.005)	—	—
Probability 3	0.500 (0.050)	0.684 (0.009)	—	0.634 (0.007)	—	—
P{ < log(0.2) }	0.052 (0.012)	0.014 (0.001)	< 0.001 ( < 0.001)	0.015 (0.001)	—	< 0.001 ( < 0.001)

P (< log(0.5))	0.149 (0.016)	0.056 (0.002)	0.068 (0.001)	0.050 (0.001)	0.061 (0.001)
P[ < log(0.8)]	0.296 (0.023)	0.144 (0.002)	0.316 (0.001)	0.138 (0.002)	0.309 (0.001)
P[ < log(0.9)]	0.367 (0.029)	0.236 (0.006)	0.410 (0.001)	0.208 (0.002)	0.407 (< 0.0001)
P (< 0)	0.442 (0.035)	0.400 (0.013)	0.500 (0.000)	0.369 (0.005)	0.500 (0.000)
P[ > log(1.111)]	0.476 (0.038)	0.386 (0.013)	0.410 (0.001)	0.384 (0.006)	0.407 (< 0.0001)
P[ > log(1.25)]	0.385 (0.039)	0.806 (0.006)	0.316 (0.001)	0.171 (0.003)	0.309 (0.001)
P[[ > log(2)]	0.149 (0.020)	0.035 (0.003)	0.068 (0.001)	0.027 (0.001)	0.061 (0.001)
P[ > log(5)]	0.038 (0.009)	0.002 (0.0003)	< 0.001 (< 0.0001)	0.001 (0.00001)	< 0.001 (< 0.0001)

*Other earnings model parameters*

$\gamma$ (lagged earnings)	0.500 (0.500)	– 0.090 (0.015)	– 0.201 (0.014)	– 0.121 (0.007)	– 0.213 (0.008)
$\tilde{\rho}$ (autocorrelation period 2)	0.500 (0.500)	0.375 (0.032)	0.529 (0.038)	0.344 (0.027)	0.398 (0.032)
$\rho$ (autocorrelation period $t$ )	0.500 (0.500)	0.652 (0.029)	0.737 (0.016)	0.655 (0.008)	0.739 (0.100)
$\phi$ (first period perm. effect)	0.400 (0.500)	0.209 (0.017)	0.225 (0.025)	0.240 (0.018)	0.320 (0.028)
$\sigma_\epsilon$ (s.d. individual shock)		0.260 (0.012)	0.215 (0.017)	0.366 (0.016)	0.302 (0.017)

*Variances and decompositions*

Disturbance variance, age 25	0.914 (0.236)	0.614 (0.045)	0.599 (0.020)	0.738 (0.057)	0.655 (0.027)
Disturbance variance, age 30	<sup>b</sup>	0.455 (0.021)	0.473 (0.013)	0.505 (0.015)	0.528 (0.010)
Disturbance variance, age 45	<sup>b</sup>	0.442 (0.020)	0.445 (0.011)	0.488 (0.014)	0.498 (0.007)
Disturbance variance, age 60	<sup>b</sup>	0.442 (0.020)	0.445 (0.011)	0.488 (0.014)	0.497 (0.008)
Fraction var. transitory, age 30	0.461 (0.290)	0.752 (0.019)	0.748 (0.019)	0.608 (0.017)	0.615 (0.013)
Fraction var. transitory, age 45	0.478 (0.299)	0.785 (0.024)	0.826 (0.025)	0.638 (0.017)	0.679 (0.017)
Fraction var. transitory, age 60	0.476 (0.299)	0.785 (0.024)	0.826 (0.025)	0.638 (0.017)	0.680 (0.017)
Correlation, ages 25 and 30	0.186 (0.271)	0.313 (0.019)	0.392 (0.018)	0.354 (0.020)	0.459 (0.023)
Correlation, ages 30 and 45	0.526 (0.305)	0.230 (0.021)	0.212 (0.023)	0.376 (0.017)	0.355 (0.015)
Correlation, ages 45 and 60	0.522 (0.316)	0.216 (0.023)	0.182 (0.023)	0.363 (0.017)	0.326 (0.016)
Correlation, ages 25 and 45	0.186 (0.257)	0.246 (0.024)	0.262 (0.028)	0.295 (0.022)	0.368 (0.030)
Correlation, ages 30 and 60	0.534 (0.291)	0.229 (0.022)	0.204 (0.024)	0.375 (0.017)	0.348 (0.016)
Correlation, ages 25 and 60	0.183 (0.258)	0.246 (0.024)	0.261 (0.029)	0.295 (0.022)	0.367 (0.031)

*Marital status period 1 covariates*

Nonwhite indicator	0.000 (0.255)	– 0.442 (0.119)	– 0.444 (0.123)	– 0.593 (0.116)	– 0.606 (0.118)
Intercept	0.000 (0.255)	2.046 (0.390)	2.049 (0.390)	4.567 (0.335)	4.549 (0.328)
Education	0.000 (0.128)	– 0.075 (0.030)	– 0.074 (0.030)	– 0.232 (0.025)	– 0.226 (0.027)

Table 3 (Continued)

<i>Marital status period t covariates</i>									
Nonwhite indicator	0.000 (0.255)	-0.771 (0.110)	-0.772 (0.113)	-0.961 (0.092)	-1.015 (0.023)				
Probit age 25, Ed = 12	0.000 (0.255)	0.028 (0.204)	0.028 (0.208)	0.078 (0.219)	0.138 (0.193)				
Probit age 40 vs. 25, Ed = 12	0.000 (0.255)	0.545 (0.124)	0.554 (0.133)	0.898 (0.116)	0.906 (0.073)				
Probit age 55 vs. 40, Ed = 12	0.000 (0.255)	-0.101 (0.236)	-0.109 (0.240)	0.294 (0.068)	0.293 (0.071)				
Probit ed 16 vs. 12, Age = 25	0.000 (0.255)	-0.207 (0.133)	-0.207 (0.132)	-0.662 (0.103)	-0.667 (0.085)				
Probit ed 16 vs. 12, Age = 40	0.000 (0.255)	0.210 (0.145)	0.218 (0.144)	-0.258 (0.076)	-0.283 (0.060)				
Probit ed 16 vs. 12, Age = 55	0.000 (0.255)	-0.066 (0.249)	-0.072 (0.250)	-0.092 (0.064)	-0.099 (0.075)				
Lagged marital status	0.680 (0.680)	0.546 (0.056)	0.536 (0.055)	0.668 (0.052)	0.645 (0.040)				
Lagged earnings	0.000 (0.510)	0.125 (0.025)	0.127 (0.025)	0.177 (0.023)	0.182 (0.227)				
<i>Other marital status model parameters</i>									
$\lambda$ (autocorrelation)	0.700 (0.700)	0.907 (0.006)	0.909 (0.006)	0.926 (0.008)	0.931 (0.002)				

<sup>a</sup>Prior moments shown are for the mixed normal model, not the normal model.

<sup>b</sup>Prior moments that do not exist.

events (e.g., lengths of spells of earnings below a specified value) conditional on various combinations of personal characteristics. Since these probabilities are based on the posterior distribution, they reflect our uncertainty about parameters as well as our uncertainty about events conditional on parameters.

All results presented here for the sample of young men are based on 10,000 iterations of the Gibbs sampler following an initial 2000 iterations which were discarded. These computations were undertaken on a Sun Model 20 workstation, and required about 25 s per iteration for each model. For the mixture model based on the full sample, all results are based on 2500 iterations of the Gibbs sampler following an initial 294 iterations which were discarded. These computations required about 332 s per iteration. For the normal model based on the full sample all results are based on 1500 iterations of the Gibbs sampler following an initial 276 iterations which were discarded. These computations required about 325 s per iteration. Computational times for the full sample are much longer than for the young men sample, because there are 48,738 rather than 15,604 person-year observations and because in the full sample 47,594 person-year observations were multiply imputed in the data augmentation step described in Appendix C, whereas this step is unnecessary in the young men sample. In all four models, the relative numerical efficiency (Geweke, 1989) of the Monte Carlo approximation ranged from about 0.05 for some posterior moments to nearly 1.0 for others. This indicates moderate serial correlation in the algorithm and is not indicative of convergence problems. Both indications are confirmed by examination of the simulated functions of interest. Multiple executions of the algorithm would provide further assurance of convergence (Gelman and Rubin, 1992) but are precluded by the computing time required.

## 5. Results

Table 3 and Figs. 1 and 2 report results for two models, mixture and normal, and two samples, young men and full. The table reports prior and posterior means and standard deviations for the parameters and some functions of interest in each model and for each sample.

### 5.1. *Earnings model, first period*

The first 10 rows of Table 3 report the results for first-period earnings. All four model/sample combinations imply that first-period earnings are substantially lower for blacks than whites, *ceteris paribus*. For example, the posterior mean for the race dummy in the mixture model based on the full sample is  $-0.195$ , implying that first-period earnings are roughly 20% lower for blacks. All four sets of results indicate that those with missing values for father's education tend

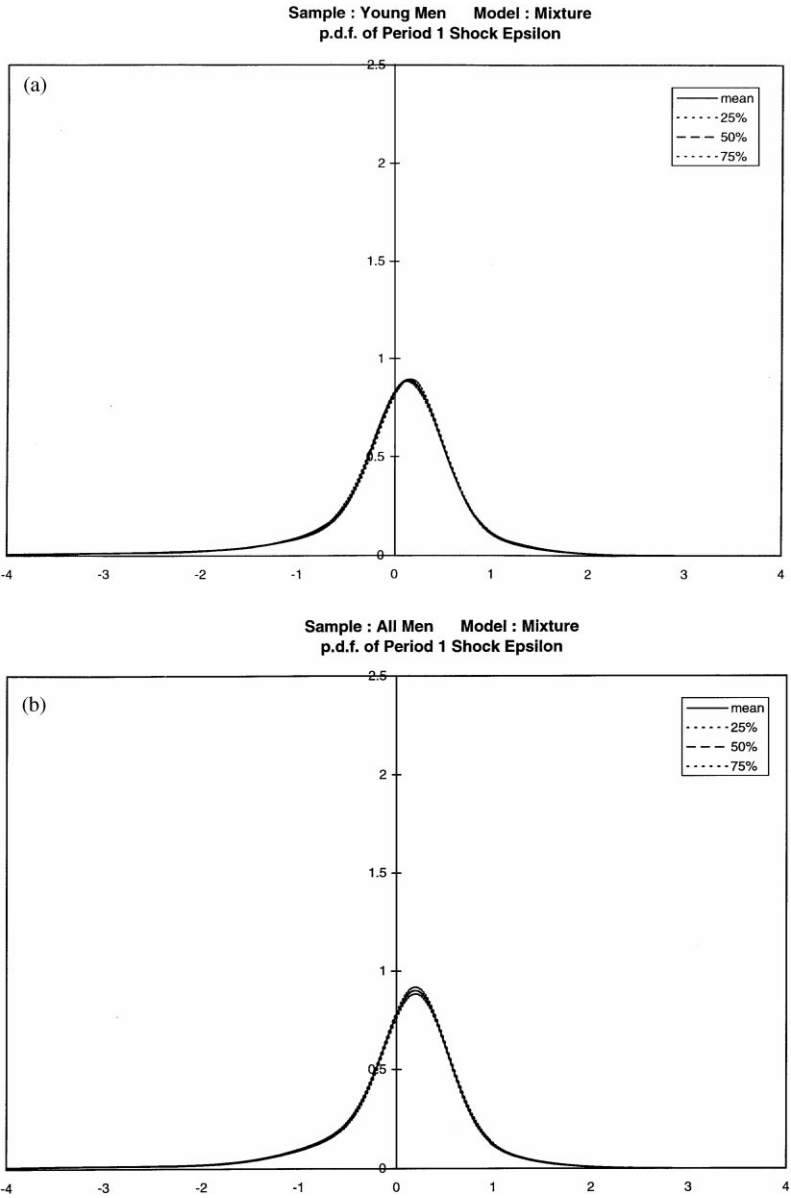


Fig. 1. Posterior distributions of p.d.f. of  $\epsilon_{i1}$ .

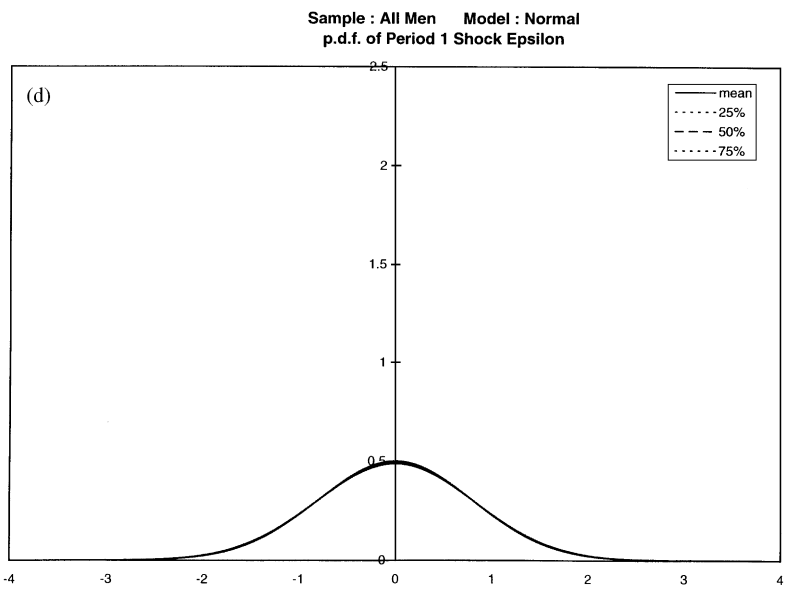
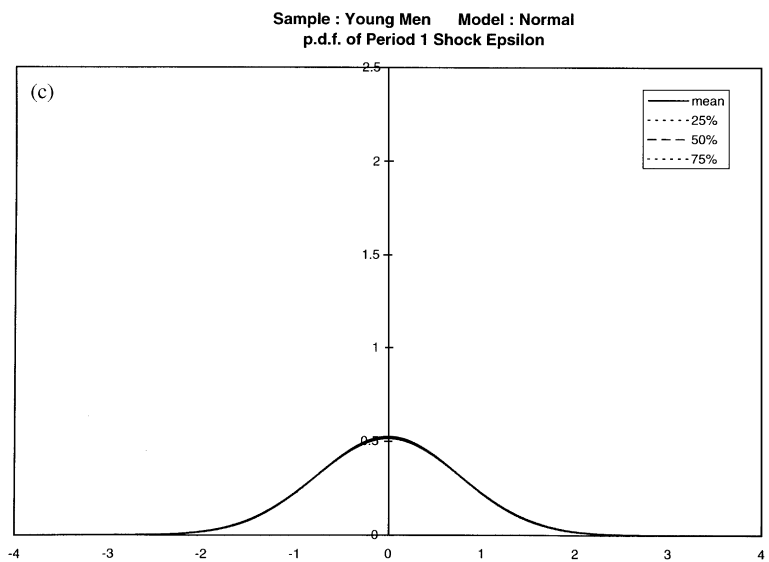


Fig. 1. (Continued)

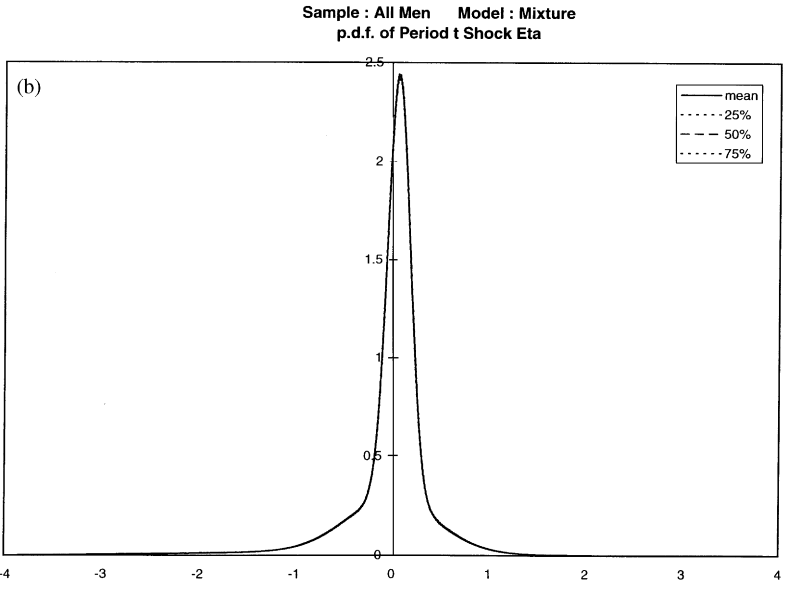
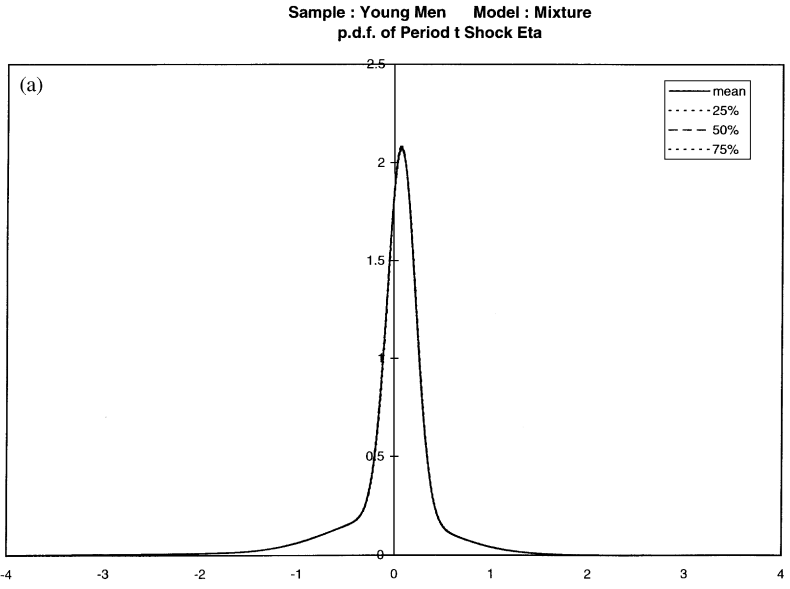
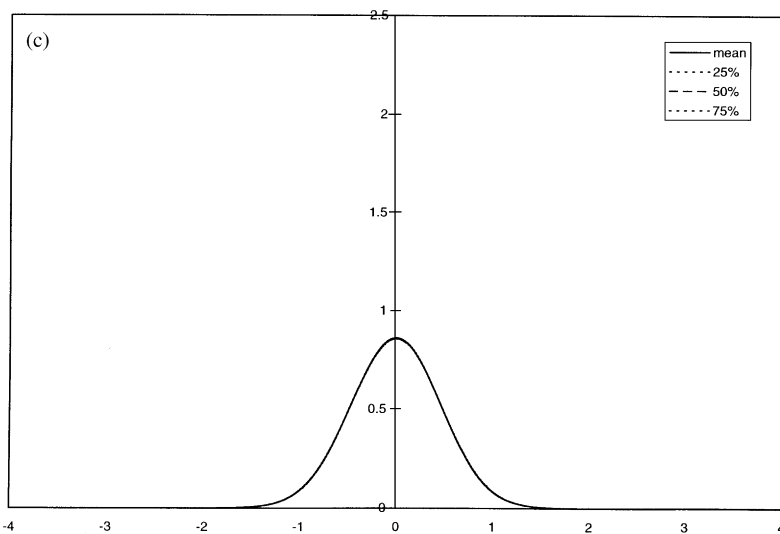


Fig. 2. Posterior distributions of p.d.f. of  $\eta_{it}$ .



Sample : Young Men Model : Normal  
p.d.f. of Period t Shock Eta



Sample : All Men Model : Normal  
p.d.f. of Period t Shock Eta

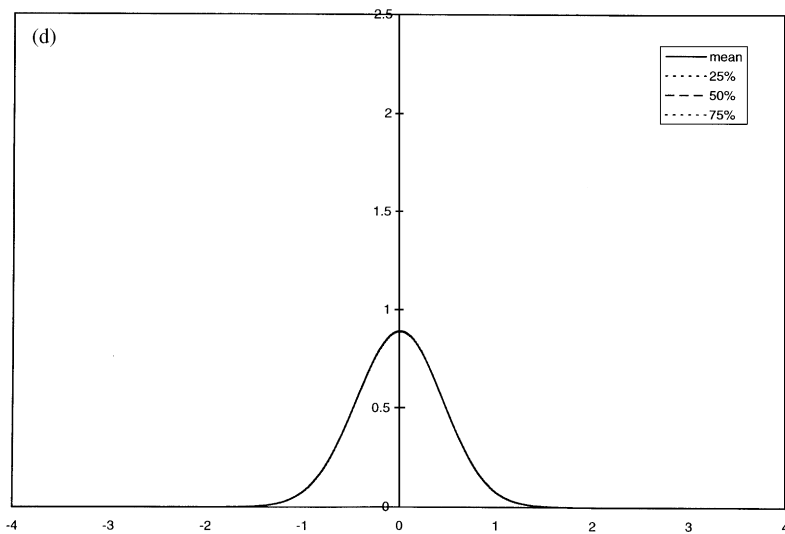


Fig. 2. (Continued)

to have lower initial earnings, but there is little evidence of any other relation between parents' education and initial earnings. In the case of parents' education the data are not very informative – comparing posterior and prior standard deviations, note that changes in prior means produce a response in posterior mean exceeding one-tenth of the prior mean change in most cases. For the race dummy the change is about one-tenth and for the other first period covariates it is less.

For the other regressors, the four sets of results imply rather different effects. For example, the mixture model based on the full sample implies that each additional year of education is associated with a 3% increase in initial earnings, while the normal model based on the full sample indicates a 12% increase. The mixture model based on the full sample provides no evidence of an association between initial marital status and initial earnings, whereas the other three models indicate that married men have initial earnings that are 7–9% greater than single men, *ceteris paribus*.

### 5.2. *Earnings model, subsequent periods*

The next 16 rows of Table 3 report the results for the model of earnings in the second period and onward. The four sets of results imply earnings ranging from 16% to 27% lower for blacks than whites, *ceteris paribus*. And the four models imply that married men have earnings that range from 4% to 10% greater than single men. The parents' education variables show no clear pattern across the models. Most are within two posterior standard deviations of zero, and posterior means respond by more than one-tenth to changes in prior means in the mixture model.

We do not report results for the parameters of the education and age polynomials, which are difficult to interpret, but rather report posterior means and standard deviations for earnings differences across certain age and education categories, corresponding to the functions  $G(a_1, a_2; e_1, e_2)$  described in Section 3.1. For example, the posterior mean for earnings at age 35 vs. 25 at education level 12 in the mixture model based on the full sample is 0.231, implying earnings growth of roughly 23% from age 25 to 35 for those with 12 yr of education. For age 45 vs. 35 the growth is 8%, whereas for 55 vs. 45 it is –4%. Thus, this model implies that earnings growth slows substantially with age and turns negative in the 50s. As another example, the posterior mean for earnings at education level 16 vs. 12 at age level 35 in the mixture model based on the full sample is 0.469, implying that college graduates earn roughly 47% more than high school graduates at age 35, *ceteris paribus*.

The coefficient  $\gamma$  on lagged earnings is negative and small in absolute value (about –0.1), whereas the serial correlation coefficient  $\rho$  is substantial (about 0.5). Thus there is strong evidence for persistence in earnings conditional on the covariates, but there is very little impact of current marital status on future earnings.

It is interesting to note that using the young men sample posterior standard deviations for the earnings at age 55 vs. 45 parameters are more than an order of magnitude greater than using the full sample. This is because in the young men sample no individual is more than 46 yr old. Thus, the data are not directly informative on earnings growth from age 45 to 55. The posterior mean for that parameter is just a combination of information from the prior and extrapolation of the age-earnings pattern from earlier ages. Notice that in the young men sample the posterior standard deviations for earnings are comparable to prior standard deviations for ages above 45, and that posterior means are all within a prior standard deviation of the prior mean at these ages. By contrast, when the sample is informative (younger ages for the young men sample and all ages for the full sample) posterior standard deviations range from 2% to 20% of prior standard deviations. This reflects the deliberate weakness of the prior (as discussed fully in Appendix 3) and the flexibility of the richly parameterized polynomial in age and education. Through this parameterization we accomplish formally what a nonparametric, non-Bayesian approach has as its informal goal: when there is no information in the data the posterior should reflect the prior, and not unwarranted extrapolation from data points with little relevance.

### 5.3. Properties of the shocks

The next two panels in Table 3 report various properties of the first-period and  $t$ th period shocks. For each shock there are 18 rows. The first nine rows report the three means, three standard deviations, and three probabilities from the mixture of three normals. Recall that the means are ordered and the second mean is set to zero, as identifying restrictions beyond the priors for these parameters (which are discussed in Appendix 3), and of course the three probabilities must sum to one: thus, there are seven free parameters. The mean of the mixture is nonzero, but since the wage equation has an intercept, the entire mixture may be renormalized to have a mean of zero. The next nine rows report some values of the cumulative distribution function (c.d.f.) for each shock, after this normalization.

Individual parameters in the mixture distribution are not tightly estimated relative to the prior. Changes in the prior means of the mixture standard deviations and probabilities often produce a corresponding change of 50% or more in the posterior mean.

Since the c.d.f.'s and probability density functions (p.d.f.'s) of these shocks are functions of the distribution parameters, posterior moments and distributions of the c.d.f.'s and p.d.f.'s are easily determined. Table 3 exhibits the c.d.f.'s at nine points, after normalization to a mean of zero. The distribution is clearly asymmetric and is very accurately determined: e.g., for the  $t$ th period shock  $\eta_{it}$ , posterior means for the full sample show the probability of a shock that cuts

wages by 50% or more is 5%, while the probability of a shock that more than doubles wages is 2.7%; posterior standard deviations are negligible.

The priors have little influence on the posterior means of the c.d.f.'s. At none of the nine points displayed in Table 3 does a change in the prior mean for the c.d.f. produce a change 10% as large in the posterior mean. This contrasts with the parameters of the mixtures themselves (first nine lines of Table 3) where the prior is more influential. Thus the data dominate in determining the posterior means of the c.d.f.'s of the shocks to earnings, but the prior is more influential in structuring the three Gaussian components of the mixture.

The implied p.d.f.'s are shown in Figs. 1 and 2. Each p.d.f. itself has a posterior distribution, reflecting uncertainty about the parameters of that distribution. To convey the p.d.f. posterior distributions, the panels plot the posterior mean, median, and quartile for each point of evaluation of the p.d.f.'s. Due to the tightness of the posterior distributions, these four are visually nearly indistinguishable. For the normal mixtures the asymmetry of the distribution is evident in every case. The mixture distributions are clearly leptokurtic, strongly skewed to the left, with modes at positive values. The mode is around  $\log(1.18)$  for the first-period shock and around  $\log(1.09)$  for the  $t$ th period shock. The normal distributions are of course symmetric. Relative to the mixture distributions they assign less probability near zero ( $\log(0.88)$  to  $\log(1.32)$  for the  $t$ th period shock), less probability far from zero (below  $\log(0.325)$  and above  $\log(4.50)$  for the  $t$ th period shock), and more probability in between.

#### 5.4. *Dynamics of the earnings model*

Of crucial importance for forecasting life-cycle earnings mobility are the covariance structure parameters and the coefficient on lagged earnings. Results for these are reported in the next 5 rows of Table 3. For example, in the mixture model based on the full sample the coefficient  $\gamma$  on lagged earnings is  $-0.121$ . This is many posterior standard deviations from zero, but small in magnitude. On the other hand, serial correlation in the shocks is substantial in magnitude,  $\rho$  having a posterior mean of 0.655. The only lagged covariate that is not perfectly collinear with the current value is marital status. Thus, the results imply that lagged marital status has very little effect on current period earnings, but there is modest serial correlation in the disturbance  $\varepsilon_{it}$  to current period earnings. The normal mixture model exhibits less serial correlation than the normal model.

In the mixture model based on the full sample, the posterior mean for the standard deviation of the individual effects is 0.366. Thus, a person with a one standard deviation above the average  $\tau_i$  value would have earnings about 37% above average, given his personal characteristics. The marital status coefficient in the same model and sample is about 0.05, implying that marital status is unimportant relative to unobserved individual effects in explaining the

variance of persistent wage differentials across men. Finally, the posterior mean for  $\phi$  in the mixture model based on the full sample is 0.240. This implies that the first-period shock could be decomposed into independent permanent and transitory components, with the permanent component having about one-fifteenth the variance of the transitory component.

Combined, these parameters imply a variance structure for disturbances to the wage equation over the lifetime. Some aspects of this variance structure are reported in the next 13 rows of Table 3. Variances are highest at age 25 but then drop quickly to a level that remains constant for the remaining years. Since all ages contain the common variance component  $(1 - \gamma)^2[\sigma_\tau^2 + \phi^2 \text{var}(\varepsilon_{i1})]$ , this is accounted for by  $\text{var}(\varepsilon_{i1}) > \text{var}(\eta_{it})/(1 - \rho^2)$ . The fraction of variance accounted for by the transitory shock  $\eta_{it}$  is about the same from age 30 onward; in the full sample this fraction is about two-thirds. Correlations between ages separated by at least five years are mainly accounted for by the permanent components  $\tau_i$  and  $\varepsilon_{i1}$ ; consistent with the fraction of variance due to the permanent component, these values are about one-third. The posterior means of all of these variances, and their decomposition, are heavily data driven. In most cases, a change in the prior mean has an impact of less than one percent on the corresponding posterior mean.

### 5.5. Marital status model

The last several rows of Table 3 contain the results for the marital status model. In the mixture model based on the full sample, the posterior mean for the education coefficient in the first-period marital status model is  $-0.232$ , implying that more educated men are much less likely to be married at age 25. But note that the posterior mean for the difference in probits between college and high school graduates at age 55 is only  $-0.092$ , implying that most of the association between education and marital status is eliminated by that age. Also interesting is that the posterior mean for the lagged earnings coefficient is  $0.177$ , implying that marital status probabilities are higher for men with greater lagged earnings.

For the young men sample, the posterior distributions of the earnings and marital status models are independent. Thus the posterior distribution of the marital status model parameters are the same in the normal and mixture models. All differences in Table 3 are due to noise in the posterior simulator. For the full sample, the posterior distributions of the two sets of parameters are linked through the unobserved earnings and marital status between age 25 and the first sample data for all men who were not in the sample at age 25. In the case of the young men sample, the posterior means and standard deviations for the marital status model in Table 3 are nearly identical across the mixture and normal models, and in the case of the full sample they are quite similar.

## 6. Simulations of earnings distributions and earnings mobility

In this section we report on simulated earnings distributions and earnings mobility for the four model/sample combinations. The simulations are performed as discussed in Section 4 and Appendix D. We first report on comparisons of simulated and actual earnings data in order to evaluate model fit. We next contrast the predictions of the four models for features of earnings distributions and earnings mobility. Finally, we compare the implications of the models for features of the distribution of the present value of lifetime earnings.

### 6.1. Model fit

Table 4 provides a comparison of the in-sample fit of the four models for the young men sample and the full sample.

As discussed in Section 2, the full sample was used to define earnings quintiles. Using these quintiles, we calculated the frequency of various quintile sequences for men in the full sample and in the young men sample. In Table 4, we use the symbol ‘–’ to denote a year in which the person is in the bottom quintile of the earnings distribution, and ‘+’ to denote a year in which he is not. For example, in the young men sample, the frequency of ‘–’ is 0.152 for whites and 0.342 for blacks. And if we look at the set of all two year sequences, the frequency of ‘– –’ is 0.089 for whites and 0.240 for blacks.

We next simulated earnings data from the four models. The simulations are based on the exogenous variables for the men in the young men and full samples (i.e. race, education, parents’ education). That is, the simulations cover only the years in which the men are observed in the respective samples, so as to allow comparison of simulated with sample earnings distributions.

Table 4 first compares the fit of the mixture and normal models to the quintile sequence data in the young men sample. The mixture model provides a much better fit to observed sequence probabilities than does the normal model. For example, the actual frequencies of ‘– – –’ and ‘+ + +’ sequences for blacks in the young men sample are 0.181 and 0.538, respectively. The mixture model predicts frequencies of 0.181 and 0.532, while the normal model predicts 0.278 and 0.355 respectively. In fact, for every sequence considered, the mixture model comes closer to replicating the sample frequency than does the normal model.

Table 4 next compares the fit of the mixture and normal models to the quintile sequence data in the full sample. With only three exceptions out of 28 cases (the ‘– + –’, ‘+ – +’ and ‘+ + –’ sequences for blacks), the mixture model comes closer to replicating the sample sequence frequencies than does the normal model. However, the agreement between sample frequencies and simulated frequencies for the mixture model is not nearly as close as it was in the young men sample. Obviously, it is more challenging to fit earnings distributions and transition frequencies for a 25–65 age range than a 25–46 age range.

Table 4  
Comparison of model fit: probabilities of earnings quintile sequences

Sequence	White			Black		
	PSID	Mixture	Normal	PSID	Mixture	Normal
<i>Sample of young men</i>						
–	0.152	0.179	0.255	0.342	0.327	0.469
+	0.848	0.821	0.745	0.658	0.673	0.531
– –	0.089	0.110	0.157	0.240	0.232	0.346
– +	0.058	0.071	0.101	0.091	0.097	0.124
+ –	0.053	0.062	0.084	0.087	0.083	0.111
+ +	0.800	0.758	0.659	0.582	0.588	0.420
– – –	0.058	0.075	0.109	0.181	0.181	0.278
– – +	0.028	0.036	0.049	0.050	0.054	0.069
– + –	0.014	0.010	0.028	0.031	0.019	0.046
– + +	0.043	0.061	0.074	0.060	0.077	0.078
+ – –	0.025	0.033	0.039	0.046	0.049	0.061
+ – +	0.026	0.029	0.045	0.040	0.033	0.049
+ + –	0.037	0.047	0.054	0.054	0.056	0.063
+ + +	0.769	0.708	0.601	0.538	0.532	0.355
<i>Full sample</i>						
–	0.144	0.206	0.249	0.348	0.422	0.491
+	0.856	0.821	0.751	0.652	0.673	0.509
– –	0.088	0.144	0.162	0.248	0.349	0.381
– +	0.044	0.060	0.083	0.084	0.076	0.110
+ –	0.051	0.057	0.081	0.089	0.072	0.105
+ +	0.817	0.739	0.674	0.578	0.504	0.404
– – –	0.061	0.112	0.118	0.196	0.306	0.318
– – +	0.020	0.031	0.040	0.042	0.045	0.064
– + –	0.012	0.011	0.026	0.033	0.019	0.043
– + +	0.031	0.048	0.057	0.051	0.057	0.067
+ – –	0.023	0.031	0.040	0.044	0.045	0.061
+ – +	0.022	0.026	0.041	0.038	0.027	0.044
+ + –	0.039	0.044	0.054	0.056	0.049	0.061
+ + +	0.792	0.696	0.625	0.539	0.452	0.343

Note: The earnings quintiles are defined using full PSID data set on male household heads aged 25–65. In the column headed ‘Sequence’, ‘–’ indicates the person is in the bottom quintile of the earnings distribution, and ‘+’ indicates they are not. A ‘– –’ indicates the person is in the bottom quintile for three consecutive years. For purposes of comparison with the PSID data, the simulation for a person covers only those periods for which data is observed for the person in the relevant PSID sample (young men or full).

It is also interesting to examine how the models fit the cross-sectional log earnings distribution at various ages. Fig. 3a reports kernel density estimates for log earnings at age 25 in both the young men sample and the simulated data from the normal model estimated with the young men sample. (This and all

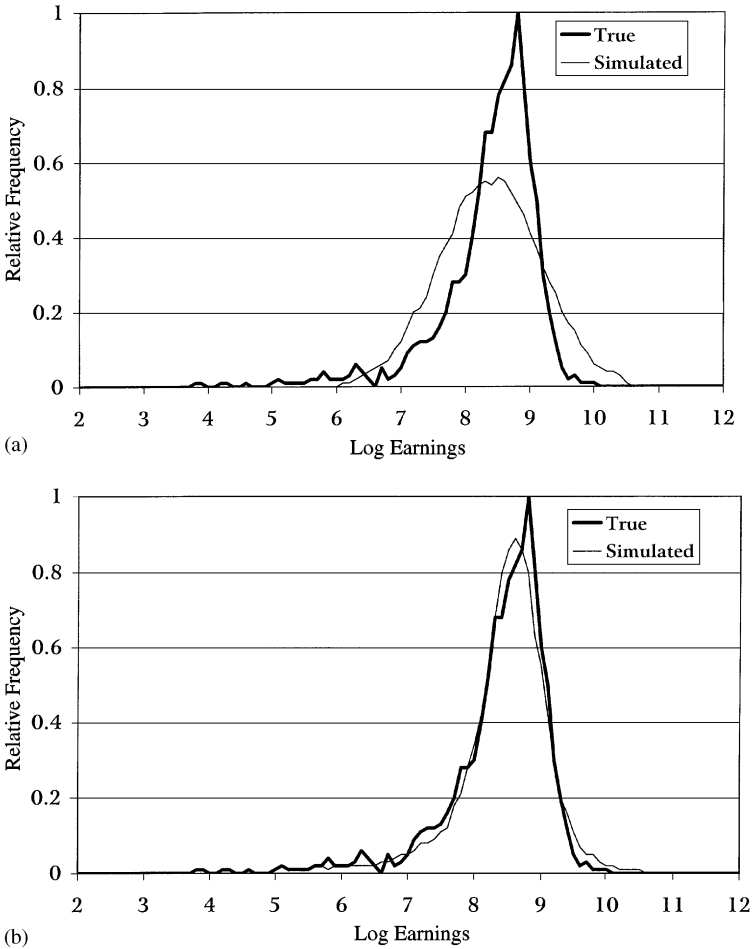


Fig. 3. Log earnings density estimate – age 25 of young samples (a) Normal (b) Normal mixture.

other density estimates reported in this paper were obtained using a Parzen kernel with a bandwidth of 0.10). As is apparent, the normal model fails to capture important features of the earnings distribution. It underestimates the mode, places too little mass near the mode, has an excessive interquartile range, and fails to capture the long left tail of the observed earnings distribution.

Fig. 3b reports the same kernel density estimates for the mixture model. Clearly, this model captures the shape of the earnings density much better than the normal model. Figs. 4a and b report similar density estimates at age 30 based on the young men sample. Again, the mixture model does much better, but not quite as well as at age 25.



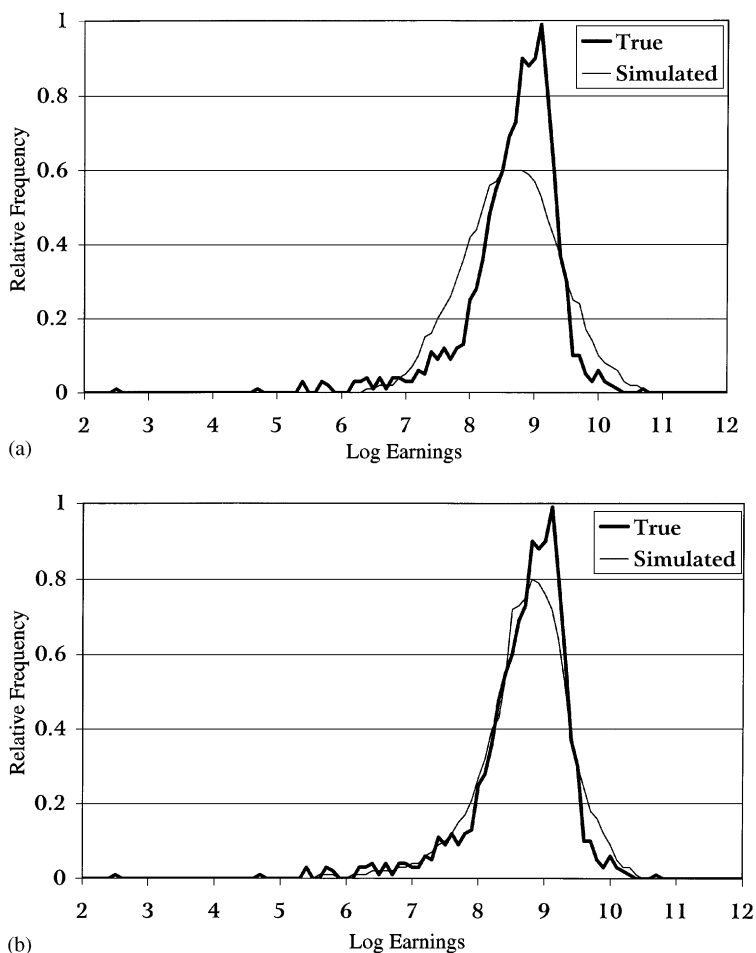


Fig. 4. Log earnings density estimate – age 30 of young samples (a) Normal (b) Normal mixture.

We now turn to evaluation of the earnings distribution in models based on the full sample. Kernel density estimates for log earnings at ages 25, 30, 45, and 60 in the full sample, using the normal model and the mixture model, are reported in Figs. 5–8. The interesting pattern in these figures is that, based on the full sample, the mixture model fits the log earnings distributions much better than the normal model at ages 25 and 30, but at age 45 it only does slightly better. As can be seen in Fig. 7b, by age 45 the mixture model suffers from the same basic set of problems that were attributed to the normal model above (i.e., it underestimates the mode, places too little mass near the mode, and has an excessive interquartile range). By age 60 the superiority of the mixture model is again

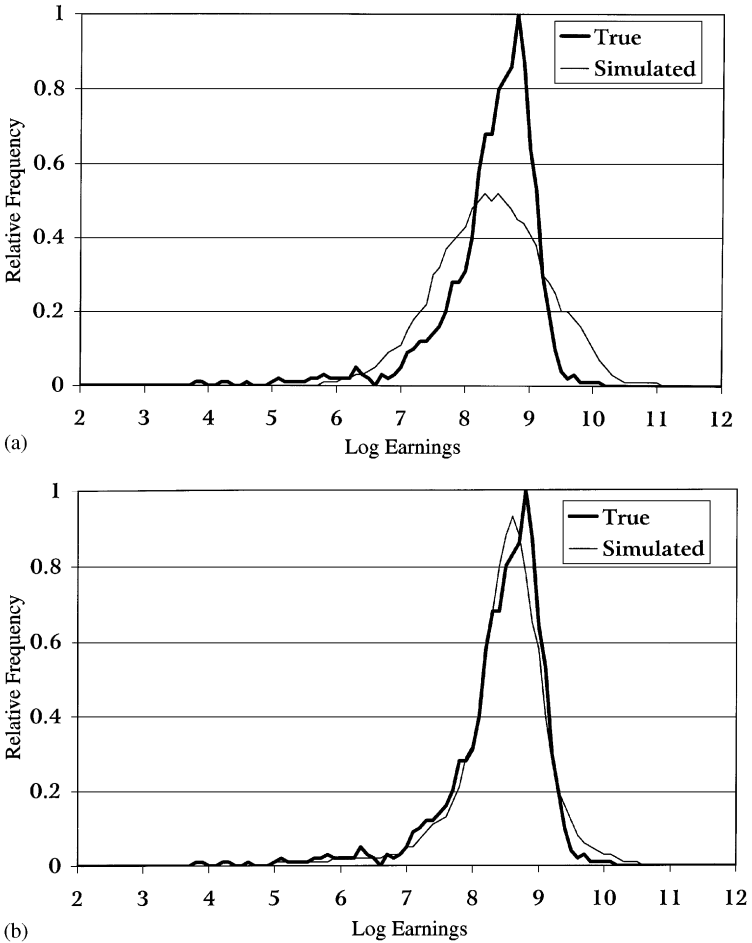


Fig. 5. Log earnings density estimate – age 25 of full samples (a) Normal (b) Normal mixture.

apparent (see Fig. 8), but discrepancies between its predictions and the observed earnings density are still obvious.

We conjecture that these problems arise because at age 45 a greater fraction of the variation in log earnings is due to variation in covariates than at either earlier or later ages. Two possible modifications of the model that may better enable us to capture the age 45 earnings distribution, and that we intend to explore in future work, are (1) to allow for a more flexible pattern of changing effects of covariates on earnings with age, and (2) to allow the variance of the shocks to vary with age (so that it may fall in the middle of the life cycle but rise at the end).

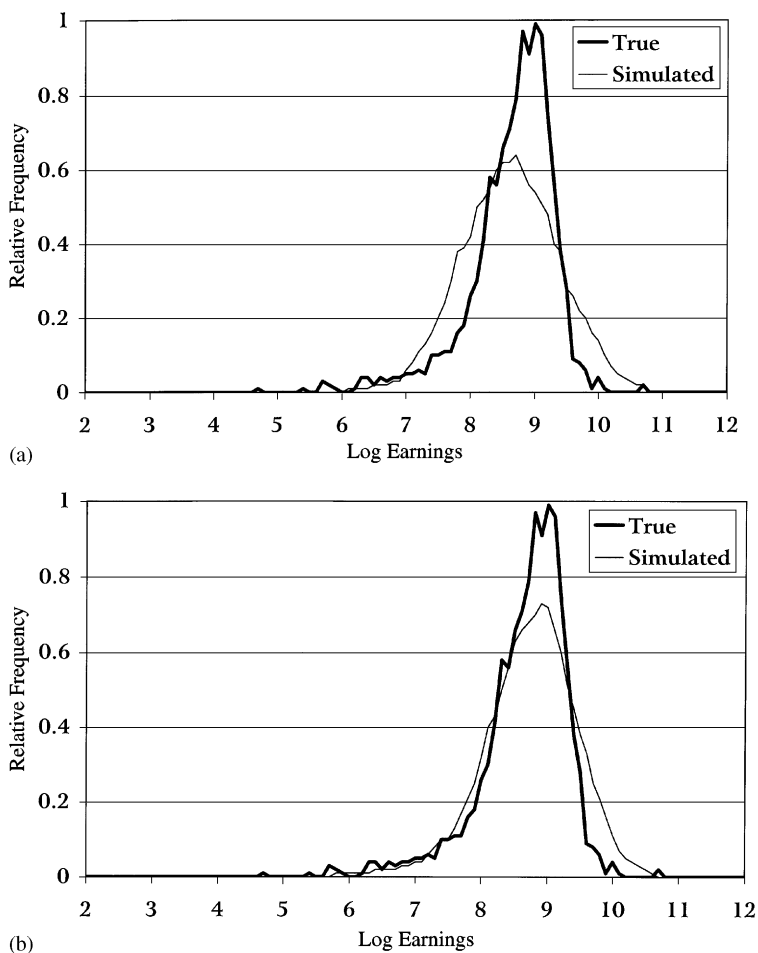


Fig. 6. Log earnings density estimate – age 30 of full samples (a) Normal (b) Normal mixture.

Finally, we explore the fit of the models to conditional log earnings distributions. Fig. 9a reports kernel density estimates for log earnings at age 35, conditioning on the event that the men were in the bottom quintile of the earnings distribution at age 34. Density estimates are reported for both the young men sample itself and the simulated data from the normal model estimated from the young men sample. (Note that only men who were observed at both ages 34 and 35 were used to generate covariates for the simulation.) As expected, the normal model places too little mass near the mode. Fig. 9b reports corresponding kernel density estimates for the mixture model. This places more mass near the mode and better captures the shape of the conditional density.

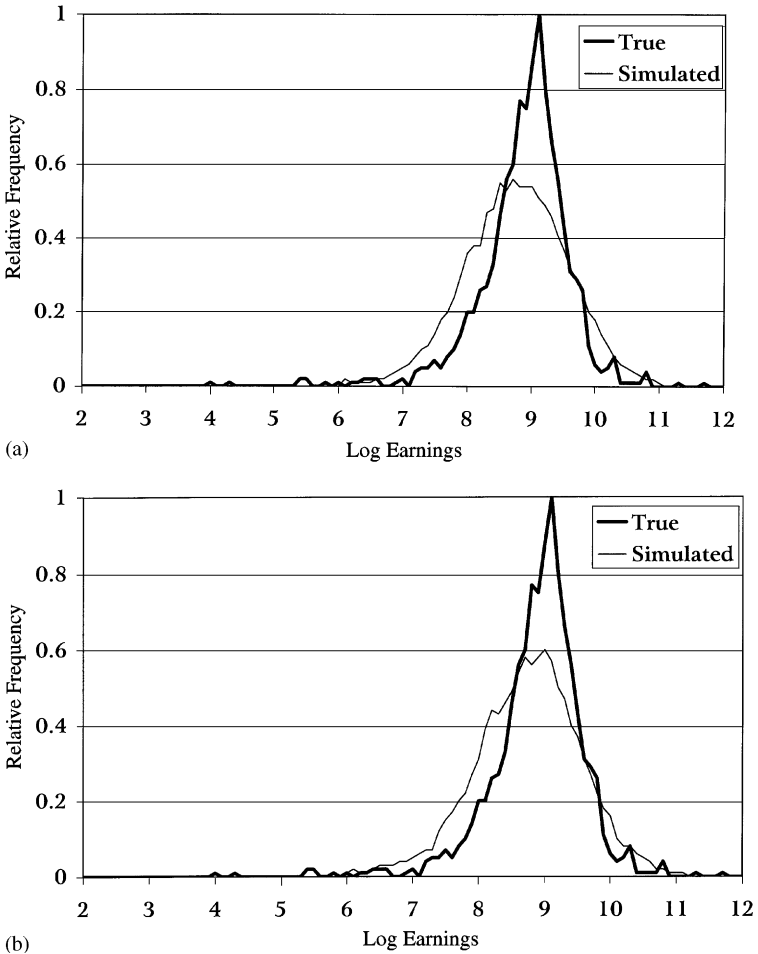


Fig. 7. Log earnings density estimate – age 45 of full samples (a) Normal (b) Normal mixture.

Fig. 10 reports similar age 35 earnings density estimates for the young men data and models, but now conditioning on not being in the bottom quintile at age 34. Comparing Figs. 10a and b, it is apparent that the mixture model fits the shape of the conditional earnings density quite closely, while the normal model does not. Figs. 11 and 12 report the same density estimates for the full sample data and models. The superior fit of the mixture model is again apparent in these figures.

To summarize, the mixture model clearly provides a better fit to log earnings distributions (both unconditional and conditional on past earnings status) in the PSID than does a normal model. However, in absolute terms the mixture model only appears to provide a close fit to these distributions at young ages (say

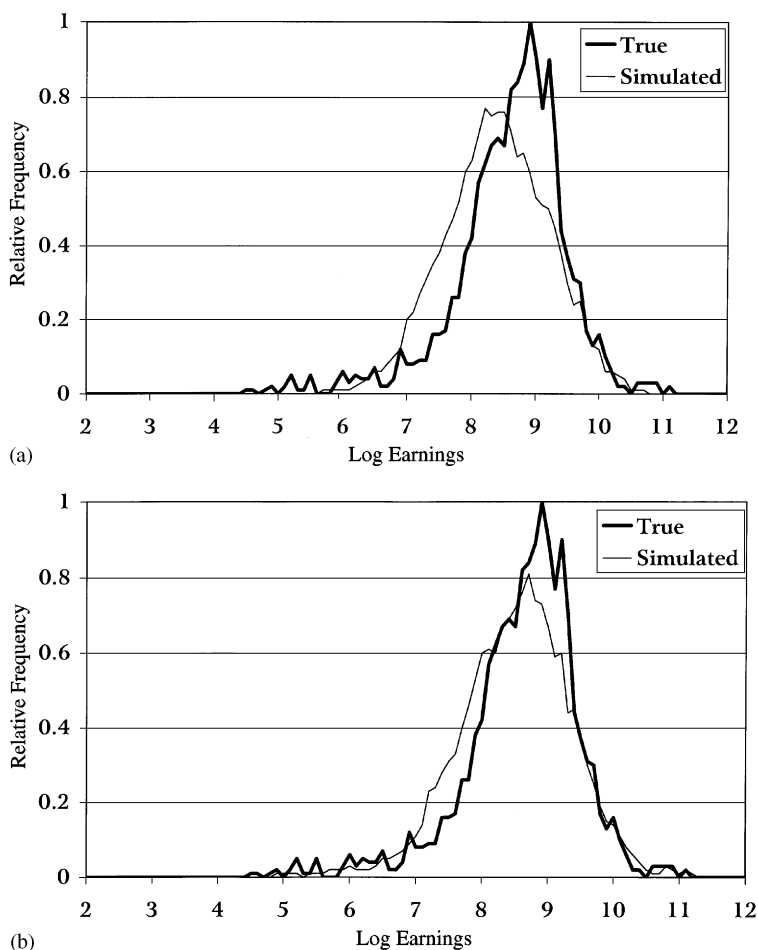


Fig. 8. Log earnings density estimate – age 60 of full samples (a) Normal (b) Normal mixture.

25–40), and again at older ages (near 60). At age 45 it fails to capture important features of the earnings distribution. We conjecture that this is due either to failure to allow for a sufficiently flexible pattern of changing effects of covariates on earnings with age, or failure to allow the variance of the shocks to vary flexibly with age. We intend to explore these issues in future research.

## 6.2. Earnings mobility

We next report on the implications of the four models for earnings mobility, focusing on mobility across quintiles. For this purpose, it is no longer necessary

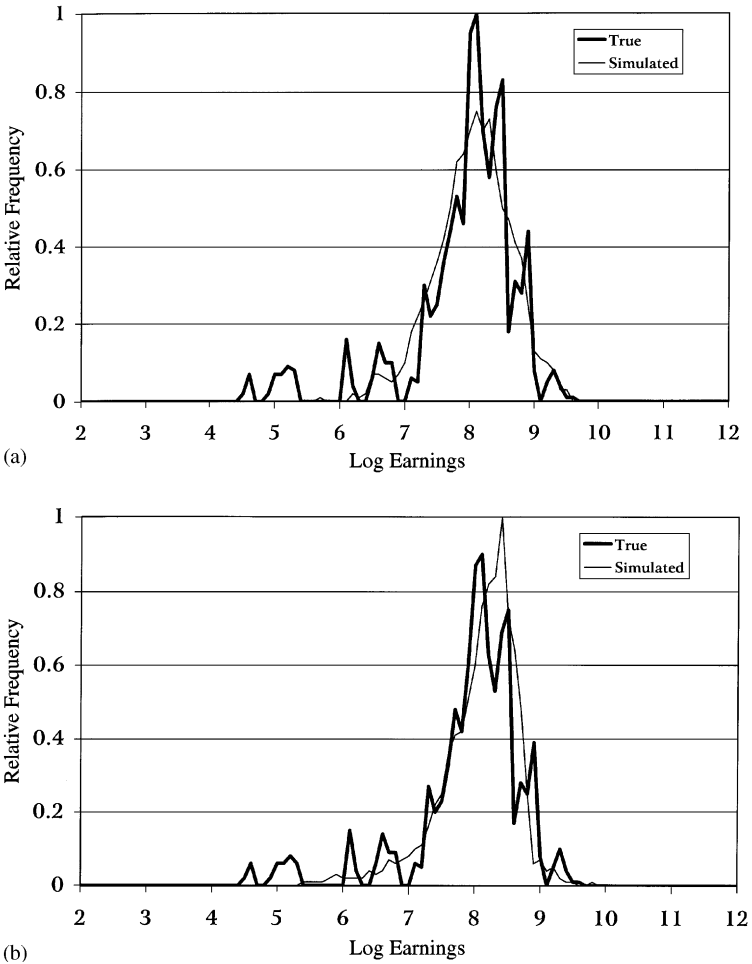


Fig. 9. Density estimate of earnings at age 35 conditional on low income at age 34 of young samples (a) Normal (b) Normal mixture.

or appropriate to restrict the simulations to the periods when a person is observed in the data. Rather, we use the same distribution of covariates as in the PSID analysis sample (i.e., race, education, parent’s education) but simulate earnings and marital status from age 25 through 65.

Tables 5a and 5b reports simulations of the probability of various earnings quintile sequences at ages 30, 45 and 60. For example, the symbol ‘— — —’ indicates that a person is in the bottom quintile at all three ages. Looking at such statistics gives one a feel for the model’s predictions regarding life cycle

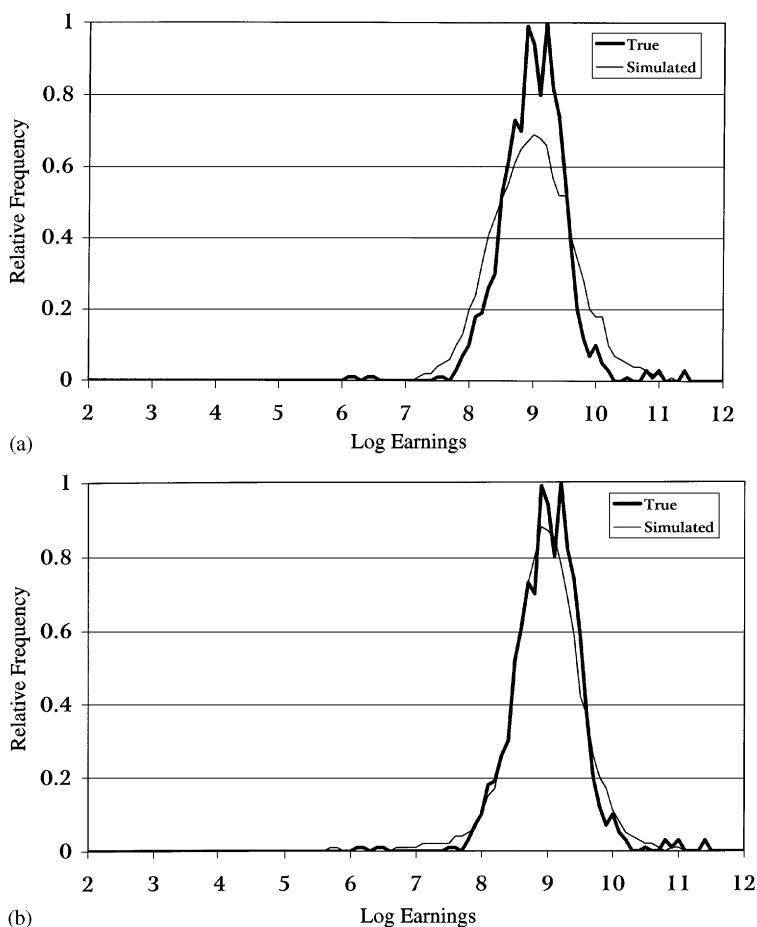


Fig. 10. Density estimate of earnings at age 35 conditional on high income at age 34 of young samples (a) Normal (b) Normal mixture.

earnings mobility. For example, in Table 5a, simulations of the mixture model based on the young men sample imply that sequences in which a male is in the bottom quintile at age 30 and then moves up (i.e., ‘ $- + +$ ’) are much more common than sequences in which a male moves into the bottom quintile at age 60 (i.e., ‘ $+ + -$ ’). For white males with 12–15 yr of education the simulated probabilities of these sequences are 0.122 and 0.047, respectively. But Table 5b, which contains results based on the full sample, shows that the same simulated probabilities in the normal mixture model are 0.102 and 0.086, respectively. Thus, the full sample results imply that falling into a low-earnings state at later ages is much more likely. These results must be given more credence, since the

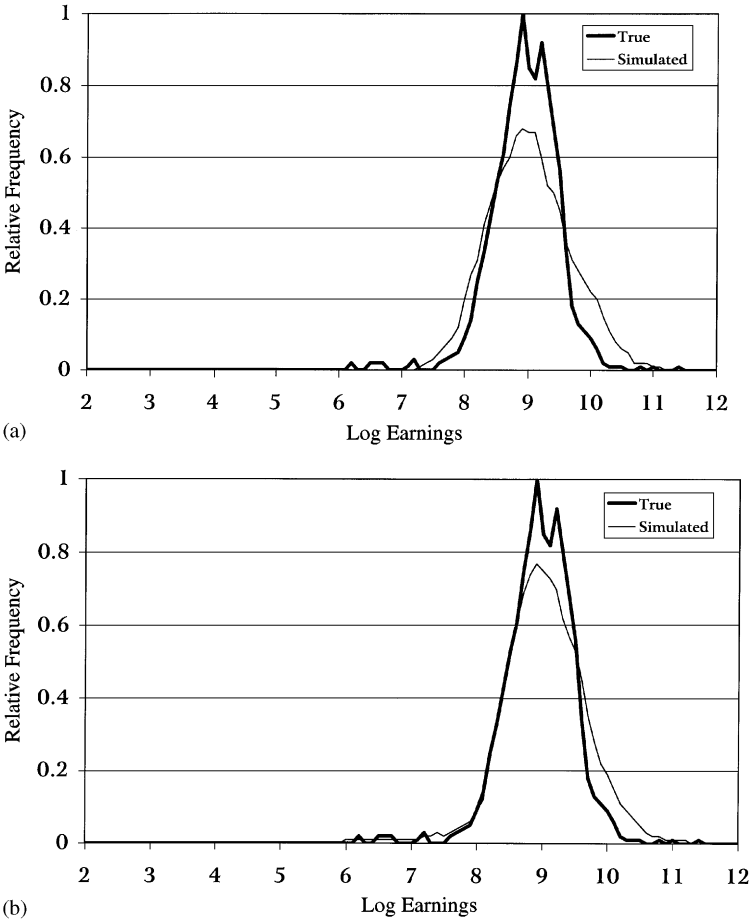


Fig. 11. Density estimate of earnings at age 35 conditional on high income at age 34 of full samples (a) Normal (b) Normal mixture.

predicted earnings changes after age 46 in the young men models are not based on data.

Table 6 reports simulations of the fraction of the lifetime (from age 25 through 65) that men in various race and education categories are predicted to spend in the bottom earnings quintile. Both expected number of spells in the bottom quintile and expected spell length are reported. The mixture model estimated from the full sample predicts that a typical white male without a high school degree will spend 47.6% of the time in the bottom quintile. His expected number of spells in the bottom quintile is 3.49, and the mean spell length is 5.45. In



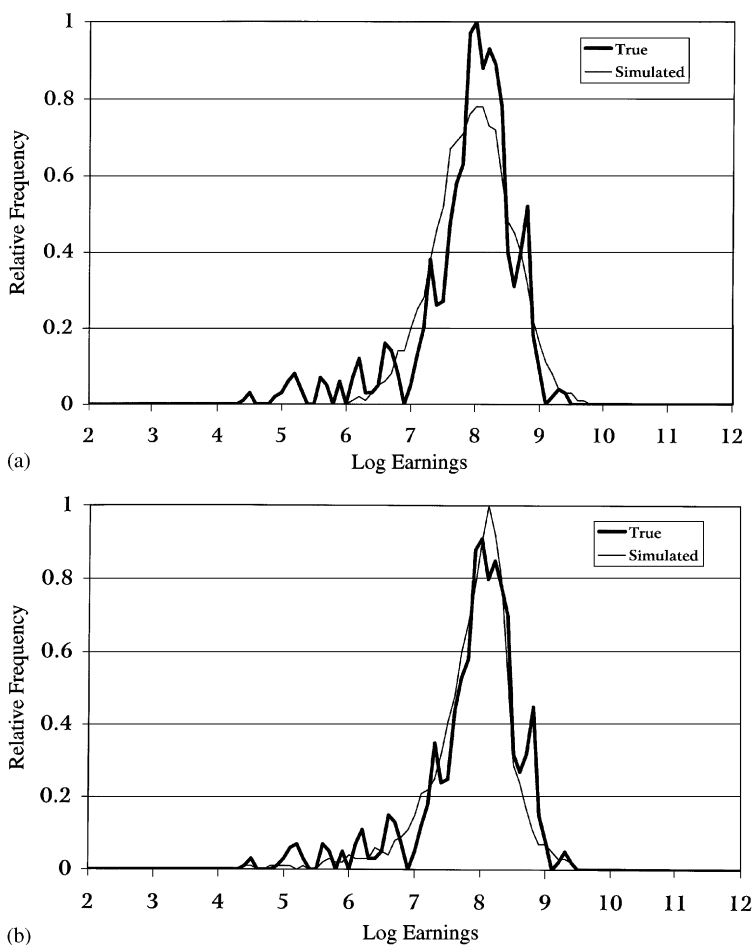


Fig. 12. Density estimate of earnings at age 35 conditional on low income at age 34 of full samples (a) Normal (b) Normal mixture.

contrast, the normal model estimated from the full sample implies a higher fraction of years in the bottom quintile (51.9%), a substantially larger expected number of spells (4.90), and a substantially shorter mean spell length (4.28 yr). Thus, the mixture model implies that low-earnings status tends to be more persistent and more concentrated among men experiencing long spells. But for white males with college degrees the mixture model implies both fewer and shorter low-earnings spells than does the normal model. This is because the mixture model simply predicts much less time in the bottom quintile for the college educated than does the normal model.

Table 5a

Simulations of earnings mobility based on young men sample, probabilities of earnings quintile sequences (ages 30, 45, 60)

Sequence	White			Black		
	Ed. < 12	Ed. 12–15	Ed. 16 +	Ed. < 12	Ed. 12–15	Ed. 16 +
<i>Mixture distribution for transitory error component</i>						
– – –	0.023	0.007	0.002	0.053	0.022	0.003
– – +	0.061	0.029	0.007	0.116	0.064	0.013
– + –	0.023	0.016	0.011	0.040	0.030	0.017
– + +	0.155	0.122	0.084	0.177	0.161	0.120
+ – –	0.017	0.007	0.003	0.018	0.012	0.010
+ – +	0.096	0.070	0.039	0.118	0.096	0.053
+ + –	0.052	0.047	0.066	0.053	0.060	0.073
+ + +	0.573	0.701	0.787	0.424	0.554	0.710
<i>Normal distribution for transitory error component</i>						
– – –	0.065	0.014	0.002	0.143	0.052	0.007
– – +	0.079	0.049	0.012	0.135	0.108	0.047
– + –	0.071	0.029	0.008	0.104	0.065	0.027
– + +	0.178	0.162	0.107	0.170	0.196	0.153
+ – –	0.043	0.013	0.003	0.051	0.034	0.013
+ – +	0.106	0.084	0.045	0.102	0.106	0.050
+ + –	0.091	0.055	0.039	0.087	0.066	0.060
+ + +	0.367	0.594	0.784	0.208	0.373	0.643

Note: Based on simulations of earnings from age 25 through 65. The sequence refers to status of being in the bottom earnings quintile at ages 30, 45 and 60. A '–' indicates the person is in the bottom quintile of the earnings distribution, and a '+' indicates he is not.

The mixture model estimated from the full sample also predicts that the fraction of time in the bottom quintile declines quite precipitously with education, especially for blacks (e.g., 60.9% when education is less than 12 yr but only 15.1% when education is 16 yr or more). The normal model predicts a less sharp decline (e.g., 68.1% for blacks with education less than 12 yr and 28.9% when education is 16 yr or more).

Table 7 reports simulations of the fraction of the remaining lifetime, from age 31 to 65, that people are expected to spend in the bottom earnings quintile conditional on low-earnings status at age 30. Based on the full sample, if a white male without a high school degree is in the bottom quintile at age 30, the mixture model predicts that he will spend 70.5% of the remaining years in the bottom quintile. If he is not in the bottom quintile at age 30, the fraction is only 29.2%. The corresponding figures for the normal model are 65.6 and 36.4%. Note that status at age 30 is predicted to have less impact on subsequent outcomes in the normal model than in the mixture model (i.e., the divergence in the two fractions is greater in the later). The same pattern holds for blacks with education less

Table 5b  
 Simulations of earnings mobility based on full sample, probabilities of earnings quintile sequences (ages 30, 45, 60)

Sequence	White			Black		
	Ed. < 12	Ed. 12–15	Ed. 16 +	Ed. < 12	Ed. 12–15	Ed. 16 +
<i>Mixture distribution for transitory error component</i>						
– – –	0.201	0.037	0.006	0.334	0.068	0.019
– – +	0.058	0.024	0.008	0.064	0.036	0.017
– + –	0.078	0.034	0.010	0.083	0.057	0.019
– + +	0.097	0.102	0.068	0.083	0.120	0.058
+ – –	0.049	0.017	0.004	0.057	0.024	0.007
+ – +	0.071	0.058	0.035	0.059	0.064	0.045
+ + –	0.088	0.086	0.046	0.081	0.097	0.062
+ + +	0.358	0.643	0.824	0.238	0.534	0.772
<i>Normal distribution for transitory error component</i>						
– – –	0.150	0.041	0.010	0.296	0.117	0.029
– – +	0.057	0.032	0.012	0.070	0.054	0.028
– + –	0.134	0.080	0.033	0.165	0.130	0.104
– + +	0.121	0.117	0.087	0.093	0.125	0.112
+ – –	0.058	0.031	0.011	0.078	0.056	0.023
+ – +	0.058	0.048	0.031	0.047	0.052	0.043
+ + –	0.129	0.140	0.097	0.109	0.142	0.119
+ + +	0.293	0.512	0.720	0.142	0.325	0.542

Note: Based on simulations of earnings from age 25 through 65. The sequence refers to status of being in the bottom earnings quintile at ages 30, 45 and 60. A '–' indicates the person is in the bottom quintile of the earnings distribution, and a '+' indicates he is not.

than 12 yr. Thus, for those with low education, the mixture model clearly predicts more persistence in low-earnings status than does the normal model.

Again based on the full sample, at higher education levels the mixture model predicts a lower fraction of remaining life in the low-earnings state than does the normal model, regardless of age 30 earnings status. This reflects the greater positive association between education and earnings in the mixture model that has already been noted.

Tables 8a and 8b report simulations of year-to-year transition probabilities at age 35. Table 8a contains results for the models based on the young men data, and Table 8b contains results based on the full sample. If a white male without a high school degree is in the bottom quintile at age 34, the full sample mixture model predicts he has an 81.2% probability of remaining in the bottom quintile at age 35. If he is not in the bottom quintile at age 34, the probability of being in the bottom quintile at age 35 is only 13.2%. The corresponding figures for the normal model are 73.4 and 17.5%. Thus, the mixture model predicts more persistence in low-earnings status.

Table 6

Simulations of earnings mobility, conditional expectations of: fraction of lifetime in bottom quintile; expected number of spells in bottom quintile; mean spell length

	White			Black		
	Ed. < 12	Ed. 12–15	Ed. 16 +	Ed. < 12	Ed. 12–15	Ed. 16 +
<i>Mixture distribution, young men sample</i>						
Fraction	0.264	0.179	0.121	0.364	0.271	0.177
# Spells	2.89	2.29	1.68	3.51	3.03	2.26
Length	3.47	2.88	2.58	4.01	3.41	2.92
<i>Normal distribution, young men sample</i>						
Fraction	0.422	0.254	0.128	0.571	0.422	0.221
# Spells	4.57	3.27	1.84	5.28	4.58	2.98
Length	3.64	2.94	2.48	4.33	3.61	2.78
<i>Mixture distribution, full sample</i>						
Fraction	0.476	0.221	0.107	0.609	0.307	0.151
# Spells	3.49	2.51	1.51	3.56	3.06	1.87
Length	5.45	3.34	2.47	6.91	3.91	2.95
<i>Normal distribution, full men sample</i>						
Fraction	0.519	0.315	0.170	0.681	0.477	0.289
# Spells	4.90	3.78	2.37	5.03	4.79	3.44
Length	4.28	3.31	2.76	5.51	4.01	3.31

Note: Based on simulations of earnings from age 25 through age 65. 'Fraction' is expected fraction of years in bottom quintile of earnings distribution. '# Spells' is expected number of spells in bottom quintile over life cycle (ending at age 65). 'Length' is mean length of low earnings spells.

It also interesting to examine the probability of being in the bottom quintile at age 35 conditional on '– +' vs. '+ –' patterns at ages 33 and 34. The mixture model predicts these probabilities are 28.0 and 62.3% for whites with education less than 12 yr, while the normal model predicts 36.7 and 55.7%, respectively, both based on the full sample. Thus, the normal mixture model predicts greater short-run persistence in the low-earnings state.

In summary, in comparing the mixture and normal models based on the full sample, we find that for men with low levels of education – the group for whom low-earnings spells are most common – the mixture models predict (1) that low-earnings spells are less frequent but of longer duration, (2) greater differences in expected fraction of the remaining lifetime in the low-earnings state depending on current low-earnings status, and (3) greater differences in the probability of low-earnings status at age  $t$  depending on low-earnings status at age  $t - 1$ . In each of these senses, the mixture model implies greater persistence of the low-earnings state than does the normal model.

Table 7

Simulations of earnings mobility, expected fraction of remaining lifetime in bottom quintile conditional on earnings status at age 30

Age 30	White			Black		
	Ed. < 12	Ed. 12–15	Ed. 16 +	Ed. < 12	Ed. 12–15	Ed. 16 +
<i>Mixture distribution, young men sample</i>						
–	0.398	0.288	0.213	0.471	0.391	0.294
+	0.190	0.134	0.095	0.249	0.190	0.138
<i>Normal distribution, young men sample</i>						
–	0.525	0.355	0.197	0.633	0.497	0.291
+	0.320	0.181	0.083	0.446	0.315	0.148
<i>Mixture distribution, full sample</i>						
–	0.705	0.437	0.250	0.796	0.524	0.387
+	0.292	0.152	0.079	0.357	0.198	0.104
<i>Normal distribution, full sample</i>						
–	0.656	0.492	0.351	0.772	0.616	0.476
+	0.364	0.225	0.123	0.491	0.339	0.192

Note: Based on simulations of earnings from age 25 through age 65. A ‘–’ indicates the person is in the bottom quintile of the earnings distribution when he is age 30, and a ‘+’ indicates he is not.

Table 8a

Simulations of earnings mobility based on young men sample, probabilities of bottom quintile at age 35 conditional on earnings status at ages 33 and 34

Age 33, 34	White			Black		
	Ed. < 12	Ed. 12–15	Ed. 16 +	Ed. < 12	Ed. 12–15	Ed. 16 +
<i>Mixture distribution for transitory error component</i>						
? –	0.646	0.588	0.409	0.722	0.649	0.571
? +	0.113	0.073	0.042	0.147	0.095	0.060
– –	0.698	0.641	0.422	0.763	0.698	0.667
– +	0.239	0.151	0.051	0.234	0.187	0.050
+ –	0.521	0.498	0.398	0.599	0.522	0.429
+ +	0.098	0.066	0.041	0.132	0.085	0.061
<i>Normal distribution for transitory error component</i>						
? –	0.707	0.580	0.409	0.747	0.684	0.477
? +	0.166	0.113	0.046	0.278	0.170	0.082
– –	0.770	0.647	0.506	0.796	0.748	0.611
– +	0.293	0.269	0.200	0.429	0.354	0.208
+ –	0.566	0.478	0.327	0.578	0.539	0.385
+ +	0.133	0.091	0.038	0.219	0.126	0.069

Note: Based on simulations of earnings from age 25 through age 65. A ‘–’ indicates the person is in the bottom quintile of the earnings distribution at age 33 or 34, and a ‘+’ indicates he is not. A ? indicates that status at age 33 is not specified.

Table 8b

Simulations of earnings mobility based on full sample, probabilities of bottom quintile at age 35 conditional on earnings status at ages 33 and 34

Age 33, 34	White			Black		
	Ed. < 12	Ed.12–15	Ed. 16 +	Ed. < 12	Ed.12–15	Ed. 16 +
<i>Mixture distribution for transitory error component</i>						
? –	0.812	0.639	0.469	0.872	0.707	0.552
? +	0.132	0.067	0.033	0.159	0.081	0.029
– –	0.854	0.702	0.547	0.906	0.765	0.639
– +	0.280	0.160	0.103	0.307	0.170	0.120
+ –	0.623	0.507	0.385	0.646	0.546	0.400
+ +	0.110	0.060	0.030	0.129	0.072	0.025
<i>Normal distribution for transitory error component</i>						
? –	0.734	0.608	0.503	0.820	0.697	0.541
? +	0.175	0.098	0.050	0.246	0.162	0.086
– –	0.792	0.689	0.615	0.862	0.753	0.569
– +	0.367	0.274	0.224	0.400	0.366	0.200
+ –	0.557	0.461	0.378	0.618	0.548	0.455
+ +	0.129	0.075	0.038	0.190	0.114	0.071

Note: Based on simulations of earnings from age 25 through age 65. A ‘–’ indicates the person is in the bottom quintile of the earnings distribution at age 33 or 34, and a ‘+’ indicates he is not. A ‘?’ indicates that status at age 33 is not specified.

### 6.3. Present value of lifetime earnings

In this section we use the various models to simulate lifetime earnings streams and construct present values of lifetime earnings. Table 9a reports present value of lifetime earnings at age 25 from simulations based on the mixture and normal models, using the young men data and an annual discount factor of 0.95. Table 9b contains the same calculations using the full sample. Note that the models based on the full sample predict a mean present value of lifetime earnings that is 13 to 15 thousand dollars (or 9–10%) less than that predicted by the models based on the young sample. This is because, as was discussed in Section 5, the young men sample contains no data on men over 46, so that its predictions for earnings growth beyond that age are based entirely on prior growth combined with extrapolations beyond age 46. These lead it to predict modest earnings growth from age 45 to 65. On the other hand, the models based on the full sample predict earnings declines at older ages. For this reason, we view models based on the full sample as providing more reliable simulations of lifetime earnings paths.

Table 9a

Present value of lifetime earnings at age 25 based on young men sample

	Mean	Stan. Dev.	Q (0.05)	Q (0.25)	Median	Q (0.75)	Q (0.95)
<i>Mixture distribution for transitory error component</i>							
All	146.5	66.8	65.0	100.3	134.2	178.8	268.0
Whites	160.3	69.4	73.2	112.6	148.7	193.9	285.9
Blacks	118.3	50.6	56.9	84.0	110.0	141.8	207.1
Whites							
Ed. < 12	127.0	52.5	60.1	91.2	118.0	153.8	222.2
Ed. 12–15	155.8	62.2	75.4	113.2	146.4	186.9	267.0
Ed. 16 +	204.0	82.4	98.8	149.2	190.7	243.7	343.5
Blacks							
Ed. < 12	103.3	41.4	49.6	76.2	96.3	124.0	181.2
Ed.12–15	122.3	50.1	60.5	89.1	114.8	146.4	206.5
Ed. 16 +	166.8	66.9	81.8	119.0	153.0	200.6	291.8
<i>Normal distribution for transitory error component</i>							
All	143.1	76.4	56.1	90.5	125.8	177.2	288.9
Whites	161.3	80.6	66.9	105.3	143.7	197.5	314.0
Blacks	106.0	49.4	47.4	71.2	95.9	129.0	198.4
Whites							
Ed. < 12	114.6	51.5	52.5	78.9	104.7	138.8	206.3
Ed.12–15	154.0	68.0	71.2	106.2	140.6	186.9	278.2
Ed. 16 +	226.0	98.7	105.2	158.3	206.8	274.0	413.6
Blacks							
Ed. < 12	87.4	38.4	41.3	60.3	80.0	105.7	157.2
Ed.12–15	110.6	46.6	51.7	77.3	101.8	134.1	198.6
Ed. 16 +	169.7	72.5	80.1	118.6	154.7	204.9	317.7

Note: Values are in thousands of 1967 dollars. Annual discount factor is 0.95. 'Q(p)' is the  $p$ th quantile.

Turning to the full sample results in Table 9b, we see that the normal and mixture models yield similar predictions for the unconditional lifetime earnings distributions. The normal model predicts roughly the same unconditional mean present value of lifetime earnings as does the mixture model (\$130,100 vs. \$131,600). They also predict similar standard deviations (\$85,400 vs. \$84,600) and similar quantile points. Kernel density estimates for the predicted distributions from the two models (not reported) are virtually indistinguishable. However, this similarity in unconditional distributions predicted by the two models masks dramatic differences in the conditional distributions they generate.

The mixture model predicts a mean present value of lifetime earnings for blacks that is \$45,100 (or 31%) less than that for whites. Much but not nearly all

Table 9b

Present value of lifetime earnings at age 25 based on full sample

	Mean	Stan. Dev.	Q (0.05)	Q (0.25)	Median	Q (0.75)	Q (0.95)
<i>Mixture distribution for transitory error component</i>							
All	131.6	85.4	38.8	74.4	111.7	165.7	290.9
Whites	145.7	89.3	45.8	84.8	125.2	183.2	315.0
Blacks	100.6	66.4	31.1	58.5	87.3	125.9	213.7
Whites							
Ed. < 12	93.3	48.7	34.0	60.0	84.0	115.9	183.4
Ed.12–15	147.4	72.4	59.3	98.1	133.5	181.7	278.9
Ed. 16 +	229.6	113.8	91.9	150.3	205.7	284.2	444.9
Blacks							
Ed. < 12	76.1	42.1	26.0	46.9	67.7	96.0	155.7
Ed.12–15	120.8	60.0	48.5	79.7	109.1	148.5	233.1
Ed. 16 +	192.9	151.1	77.1	122.7	171.9	238.5	353.7
<i>Normal distribution for transitory error component</i>							
All	130.1	84.6	39.9	72.9	109.4	163.9	290.7
Whites	146.3	89.8	48.2	85.0	125.0	183.1	315.8
Blacks	94.4	57.5	31.4	55.0	81.1	117.9	200.2
Whites							
Ed. < 12	102.4	56.4	37.6	63.5	90.0	127.2	208.2
Ed.12–15	149.3	79.7	58.1	94.0	131.8	184.6	301.7
Ed. 16 +	212.5	114.8	83.1	133.4	187.5	263.5	420.2
Blacks							
Ed. < 12	75.7	42.6	26.8	45.8	66.2	95.1	158.6
Ed.12–15	110.3	58.6	43.4	69.4	97.7	136.6	221.3
Ed. 16 +	160.7	91.2	59.1	97.5	140.7	193.1	332.6

Note: Values are in thousands of 1967 dollars. Annual discount factor is 0.95. 'Q(p)' is the  $p$ th quantile.

of this difference is accounted for by the education differences between blacks and whites. For example, among those with education in the 12–15 yr range, the mean present value of lifetime earnings for blacks is predicted to be \$26,600 (or 18%) less than that for whites.

The normal model predicts a greater unconditional black-white differential (\$51,900 or 35.5%), and greater black-white differentials within education classes. For example, among those with education in the 12–15 yr range, the mean present value of lifetime earnings for blacks is predicted to be \$39,000 (or 26%) less than that for whites.

The mixture model predicts greater differences in earnings across education classes than does the normal model. For example, comparing whites in the 16 +



vs. 12–15 yr of education categories, it predicts a \$82,200 (or 56%) greater present value of lifetime earnings for members of the college-educated group. The normal model predicts only \$63,200 (or 42%) greater earnings for members of the college-educated group.

The mixture model predicts that the increase in mean present value of lifetime earnings in moving from the 12–15 to 16 + yr of education group exceeds the increase in the median (e.g., \$82,200 vs. \$72,200 for whites). The univariate normal model also predicts that the increase in the mean is greater than the increase in the median (e.g., \$63,200 vs. \$55,700). Thus, both models predict that at higher education levels the present value of lifetime earnings distribution becomes more skewed to the right. This is apparent in Figs. 13 and 14, which

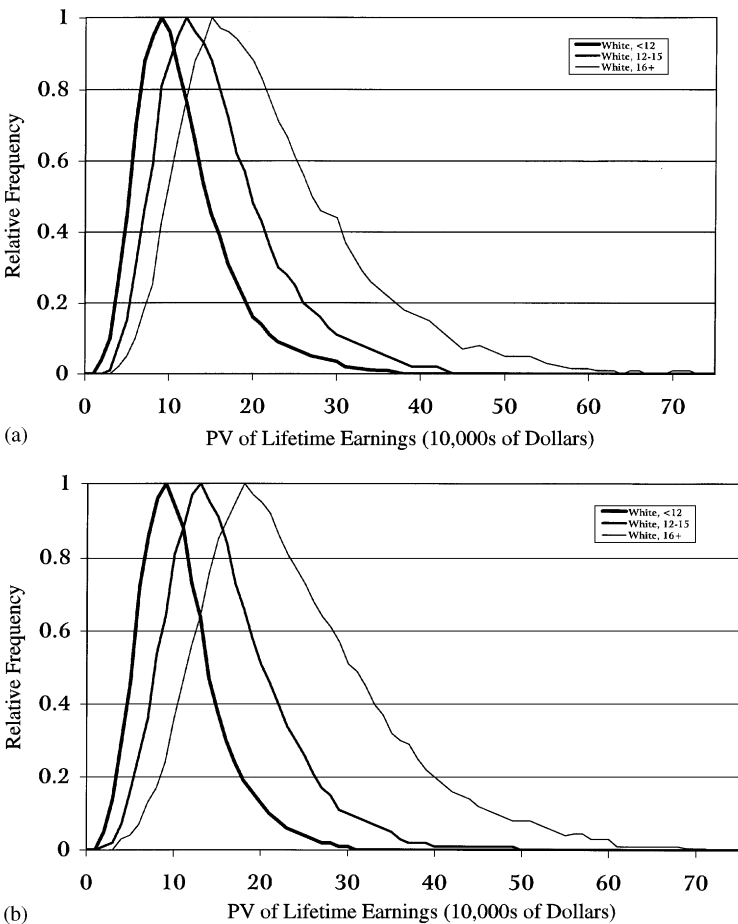


Fig. 13. Lifetime earnings density estimate by type, full samples (a) Normal (b) Normal mixture.

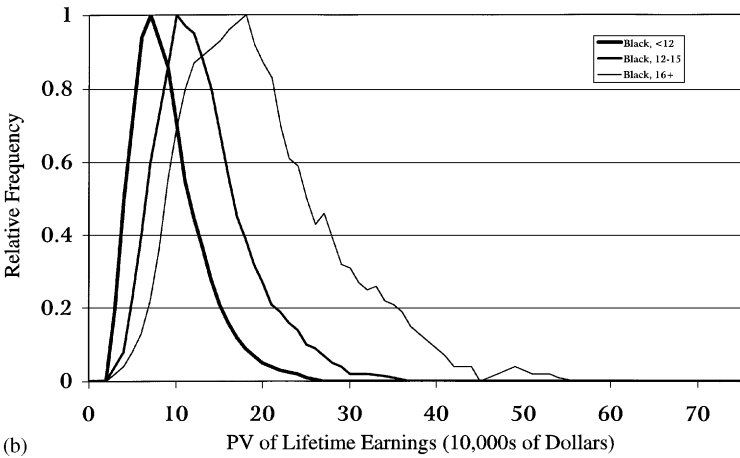
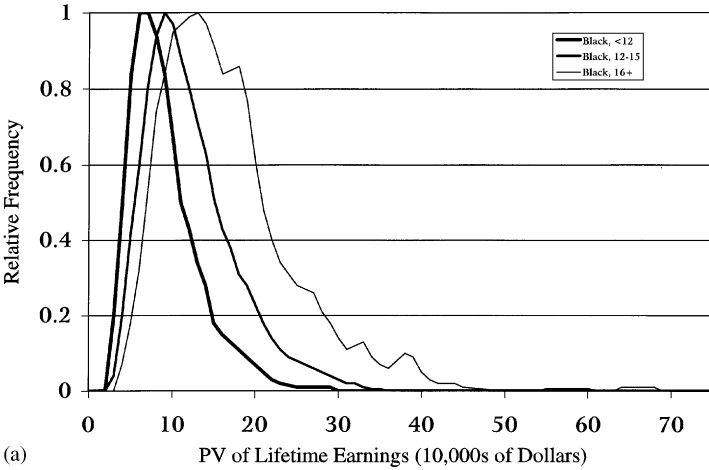


Fig. 14. Lifetime earnings density estimate by type, full samples (a) Normal (b) Normal mixture.

report kernel density estimates for the earnings distributions within race and education classes. Fig. 13a reports results for whites based on the normal model, while Fig. 13b reports results for whites based on the mixture model, both using the full sample. Comparison of Figs. 13a and b shows clearly that the mixture model predicts a greater shift right in the mode with increasing education than does the normal model. Fig. 14 shows that this difference is even more apparent for blacks.

Tables 10a and 10b report the distributions of lifetime earnings over ages 31–65, conditional on earnings quintile status at age 30, using the full sample

Table 10a

Present value of lifetime earnings, ages 31–65 based on full sample, normal distribution for transitory error component

	Mean	Stan. Dev.	Q (0.05)	Q (0.25)	Median	Q (0.75)	Q (0.95)
<i>Conditional on being in bottom quintile at Age 30</i>							
All	80.2	40.3	30.9	51.7	72.2	99.9	155.3
Whites	89.7	42.7	37.0	59.7	81.8	110.1	169.1
Blacks	67.9	33.2	27.0	44.2	61.4	85.0	130.6
Whites							
Ed. < 12	74.8	34.0	32.3	50.9	68.8	92.4	138.9
Ed.12–15	98.7	41.4	46.2	69.6	91.3	119.0	178.3
Ed. 16 +	126.8	55.2	59.1	89.3	117.2	152.0	230.6
Blacks							
Ed. < 12	59.5	28.8	24.2	39.2	53.9	73.5	114.1
Ed.12–15	80.7	34.4	37.1	56.4	73.8	99.1	146.3
Ed. 16 +	101.7	41.7	49.3	71.8	91.6	123.1	184.7
<i>Conditional on not being in bottom quintile at Age 30</i>							
All	161.6	88.4	63.8	101.5	140.9	197.8	328.7
Whites	173.2	92.1	70.6	110.6	151.9	211.4	346.2
Blacks	124.6	62.4	54.0	82.3	110.6	152.3	240.9
Whites							
Ed. < 12	127.2	58.6	57.3	86.4	115.6	154.6	235.5
Ed.12–15	169.8	80.5	76.3	114.3	152.7	206.0	321.8
Ed. 16 +	229.8	113.3	99.0	150.1	205.2	279.2	443.2
Blacks							
Ed. < 12	102.7	45.7	47.0	70.5	93.5	123.8	187.0
Ed.12–15	133.9	61.3	61.8	91.0	120.8	161.7	250.4
Ed. 16 +	185.6	91.0	84.1	126.3	166.9	223.0	354.5

Note: Values are in thousands of 1967 dollars. Annual discount factor is 0.95. 'Q(p)' is the  $p$ th quantile.

results; Table 10a is based on the normal model and Table 10b is based on the mixture model. The mixture model implies that mean present value of lifetime earnings over ages 31–65 is \$159,100 if a male is not in the bottom quintile at age 30, and \$70,700 if he is in the bottom quintile. This is \$88,400, or 56% less. Interestingly, conditioning on race and education class does little to eliminate this difference. For example, for whites in the 12–15 yr of education range, the difference is \$161,900 vs. \$92,500. This is still \$69,400, or 43% less. Thus, even conditioning on race and education status, position in the earnings distribution at the single age of 30 reveals a great deal about expected future earnings.

It is also interesting to compare the normal and mixture model implications for lifetime earnings. According to Table 10a, the normal model implies a mean

Table 10b

Present value of lifetime earnings, ages 31–65 based on full sample, mixture distribution for transitory error component

	Mean	Stan. Dev.	Q (0.05)	Q (0.25)	Median	Q (0.75)	Q (0.95)
<i>Conditional on being in bottom quintile at Age 30</i>							
All	70.7	42.0	25.0	43.5	61.4	86.1	147.5
Whites	78.6	46.7	28.5	48.4	67.7	95.8	164.3
Blacks	60.4	31.9	22.1	38.5	54.3	74.9	120.7
Whites							
Ed. < 12	60.8	29.0	25.5	40.8	56.4	74.7	112.3
Ed.12–15	92.5	45.5	39.5	61.6	83.2	111.2	178.8
Ed.16 +	142.1	77.0	53.0	87.9	125.0	176.7	284.5
Blacks							
Ed. < 12	52.0	24.9	20.5	34.5	47.8	63.8	98.1
Ed.12–15	79.7	36.6	35.3	55.2	72.2	96.7	147.6
Ed.16 +	99.9	47.4	39.6	73.5	92.3	115.6	188.1
<i>Conditional on not being in bottom quintile at Age 30</i>							
All	159.1	92.8	64.9	98.8	136.1	193.9	331.0
Whites	169.8	94.7	69.0	105.6	145.9	207.1	350.7
Blacks	128.7	79.9	57.5	84.8	113.2	153.9	250.6
Whites							
Ed. < 12	112.6	49.2	55.3	79.0	101.9	134.0	205.7
Ed.12–15	161.9	73.4	77.1	111.1	145.8	195.6	295.2
Ed.16 +	248.5	119.8	108.5	164.4	222.3	304.7	477.6
Blacks							
Ed. < 12	101.0	43.7	49.8	71.8	92.4	119.9	180.7
Ed.12–15	138.0	61.7	68.0	95.6	124.6	164.7	252.3
Ed.16 +	214.9	196.1	95.3	138.6	190.0	257.1	387.9

Note: Values are in thousands of 1967 dollars. Annual discount factor is 0.95. 'Q(p)' is the *p*th quantile.

present value of earnings over ages 31–65 of \$161,100 if a man is not in the bottom quintile at age 30, and \$80,200 if he is in the bottom quintile. This is \$80,900, or 50% less, which is less than the difference predicted by the mixture model.

Table 11 presents regressions of present values of lifetime earnings on individual characteristics, using the simulated data from the four models. The estimated partial effects of covariates are quite different across the models. For example, using the full sample the mixture model predicts that blacks have an

Table 11

Present value of lifetime earnings at Age 25; Regression on race, education, parents' education,  $\tau_i$ ,  $\varepsilon_{i1}$ 

	Young men sample		Full sample	
	Mixture model	Normal model	Mixture model	Normal model
Education	11.282 (0.237)	15.248 (0.255)	12.985 (0.108)	10.278 (0.113)
Black	– 28.497 (1.031)	– 35.305 (1.111)	– 19.214 (0.720)	– 30.724 (0.752)
Father's Ed.				
Missing	– 22.454 (2.384)	– 14.910 (2.566)	– 4.472 (1.593)	– 8.458 (1.663)
HS	– 0.477 (1.057)	7.408 (1.138)	13.607 (0.834)	11.209 (0.871)
College	1.461 (1.713)	5.678 (1.844)	– 1.233 (1.414)	4.120 (1.476)
Mother's Ed.				
Missing	10.372 (2.881)	10.110 (3.102)	– 8.989 (0.975)	– 6.534 (1.018)
HS	– 2.149 (0.943)	– 2.808 (1.015)	– 2.682 (0.714)	– 2.705 (0.745)
College	– 12.138 (2.190)	– 5.668 (2.358)	5.756 (1.589)	7.780 (1.660)
$\tau_i$	139.224 (1.137)	134.483 (1.573)	126.337 (0.561)	124.374 (0.690)
$\varepsilon_{i1}$	34.301 (0.383)	45.146 (0.436)	30.954 (0.245)	52.113 (0.259)
Constant	17.224 (2.993)	– 37.521 (3.222)	– 12.586 (1.361)	20.548 (1.421)
$R^2$ -Regressors	0.2115	0.3019	0.3367	0.2630
$R^2$ - $\tau_i, \varepsilon_{i1}$	0.4492	0.3565	0.3866	0.4457
Unexplained	0.3393	0.3416	0.2767	0.2913
Dependent variable:				
Mean	146.501	143.139	131.586	130.066
SD	66.810	76.435	85.413	84.607

Note: The dependent variable is the present value of lifetime earnings from a simulated lifetime earnings path, using a discount factor of 0.95 from ages 25 to 65. Earnings are measured in thousands of 1967 dollars. The number of simulated paths is 10 times the number of individuals observed in the PSID data. The distribution of education, race and parents' education is the same as in the data.

expected present value of lifetime earnings that is \$19,214 less than whites, *ceteris paribus*. This is 14.6% of the unconditional mean present value of earnings of \$131,586 in that model. But, using the same sample, the normal model predicts a black-white differential of \$30,724, which is 23.6% of the mean.

Using the full sample, the mixture model predicts that each additional year of education is associated with a \$12,985 increase in expected present value of lifetime earnings, as compared to \$10,278 in the normal model. Thus, the mixture model predicts that, *ceteris paribus*, a college graduate will have a present value of earnings that is \$51,940 greater than that for a high school graduate (39.5% of the mean). The normal model predicts a college premium of only \$41,112 (31.6% of the mean).

Largely due to the greater earnings differences across education classes predicted by the mixture model, the covariates explain 33.7% of the variation in present value of earnings in that model, while in the univariate normal model they explain only 26.3%. The mixture model attributes 38.7% of the variation in lifetime earnings to unobserved individual effects ( $\tau_i$  and  $\varepsilon_{i1}$ ) while the univariate normal model attributes 44.6%. The fact that a larger percentage of variance is due to unobserved heterogeneity than due to race, education, and family background explains why, in Tables 10a and 10b, earnings status at the single age of 30 was so important in predicting future earnings, even after conditioning on race and education.

## 7. Conclusion

In this study we have used data from the PSID on male household heads to address a number of questions about life-cycle earnings mobility. We developed a dynamic, reduced-form model of earnings and marital status, and applied it using most of the available male household data in the PSID. The model developed is nonstationary over the life cycle and permits possibly non-Gaussian shocks. Both to facilitate the use of most of the sample and to infer properties of life-cycle earnings accounting for parameter uncertainty, we developed a Gibbs sampling-data augmentation algorithm that provides numerical approximations to the exact posterior distribution of properties of earnings paths. An important property of this algorithm is that it copes with the potentially very complex distribution of earnings and marital status observations that are available for short segments of an individual's work history that do not include the initial period.

We reached several firm conclusions about life-cycle earnings mobility. First, we found that Gaussian shocks did not account for observed transition patterns between low-earnings states and higher-earnings states observed in the PSID for male household heads. The poor fit provided was similar to that reported in previous efforts to model these transitions. When non-Gaussian shocks were permitted, we found that our model accounted very well for transitions of young men in and out of low-earnings states, and it performed almost as well for older men. For men in their forties, the fit to observed transitions was not as good, but still substantially better than for the Gaussian version of the model and in comparison with previous studies.

We reported specific transition probabilities in and out of low-earnings states, exhibiting variations over race and education classifications. One of our central findings is that low earnings at a specific age, like 30, is a strong predictor of low earnings later in life, even conditioning on race, education, and age. Our model decomposes earnings into permanent and transitory components. Posterior distributions of these components show that in a *given* year, 60% to 70% of the

variation in the *log* of earnings not explained by covariates is accounted for by transitory components whose serial correlation is relatively weak, about 0.7 from year to year. But over a lifetime transitory components (by definition) average out. The posterior distributions show that about 60% of the variation of lifetime earnings that is not explained by education and race is attributable to permanent individual characteristics that are unobserved and uncorrelated with education, age, and race. This is consistent with the explanatory power of low earnings early in life for lower earnings later in life.

The non-Gaussian shock distribution has important consequences for lifetime earnings. Using the full sample of male household heads aged 25–65 in the PSID, the mixture model implies that, *ceteris paribus*, the present value of lifetime earnings is \$253,467 greater for those with a college degree than for those with only a high school degree (in 1998 dollars). The normal model implies a college premium of only \$200,626. And while the normal model predicts that present value of lifetime earnings is \$149,933 less for blacks in 1998 dollars (*ceteris paribus*), the mixture model predicts a substantially smaller black/white differential of \$93,764. Thus, the mixture model implies a greater association between education and earnings and a lesser association between race and earnings than does the normal model. The mixture model also predicts, among men with low levels of education, more persistence in low-income states than does the univariate normal model.

Although this work is in many respects complete, the model might be improved in a number of ways without radical modification. We plan to experiment further with functional forms for the age and education covariates and with the introduction of age-specific heteroscedasticity in an effort to account for dynamic mobility in middle age as well as the model now does for younger and older men. Our experience with non-Gaussian distributions was much more successful than we had hoped, both in the ability of the model and data to yield precise posterior information about these distributions and in the distributions themselves to account for earnings mobility. We therefore plan to experiment with higher-order normal mixture distributions for both transitory and permanent disturbance components.

Finally, one of our key findings is that the mixture model predicts a greater association between low earnings at a specific age (say, age 30) and the present value of earnings over the rest of life (after conditioning on race and education) than does the Gaussian model. This means that the two models would have rather different implications for how an increase in the variance of transitory earnings shocks would alter the dispersion in the present value of lifetime earnings. We thus conclude that the development of a good model of earnings mobility is an important first step toward beginning to assess how the widely noted increase in the cross-sectional variance of earnings that has occurred in the U.S. in recent decades will translate into increased dispersion in the present value of lifetime earnings. We plan to explore this question in future research.

**Acknowledgements**

We thank Garrett TeSelle and Susumu Imai for assistance in preparation of the data, and Dan Houser and Lance Schibilla for additional research assistance. Support for this work has been provided in part by a grant to the Institute for Research on Poverty, University of Wisconsin – Madison, from the Office of Assistant Secretary for Planning and Evaluation in the U.S. Department of Health and Human Services and by grants from the National Science Foundation. The views expressed in this paper are not necessarily those of the sponsoring institutions.

**Appendix A. Posterior distribution for the earnings model**

The panel consists of  $n$  individuals,  $i = 1, \dots, n$ . In this appendix it is assumed that individual  $i$  is observed in periods  $1, \dots, T_i$ ;  $T_i \leq T$  ( $i = 1, \dots, n$ ). Let  $\Omega_j = \{i: T_i \geq j\}$ , the set of individuals observed in period  $j$ , and let  $N_j$  denote the cardinality of  $\Omega_j$ . The total number of observations is  $N = \sum_{i=1}^n T_i = \sum_{t=1}^T N_t$ .

The shocks in the earnings model are distributed as mixtures of normal distributions. That is, a generic shock  $\varepsilon_i$  has distribution  $\varepsilon_i \sim N(\alpha_i, h_i^{-1})$ , and the distribution of  $\alpha_i$  and  $h_i$  are specified in a separate model. In our application the shocks are discrete mixtures of normals. These distributions, including priors, are discussed in Appendix F, which details complete conditional distributions for the  $\alpha_i$  and  $h_i$ .

For ( $t = 1, \dots, T_i$ ;  $i = 1, \dots, n$ ), let

$y_{it}$  = log real earnings of individual  $i$  in period  $t$ ,

$\mathbf{x}'_{i1}$  =  $k_1 \times 1$  vector of contemporaneous first period variables for individual  $i$ ,

$\mathbf{x}'_{it}$  =  $k_2 \times 1$  vector of contemporaneous variables and other variables lagged one period for individual  $i$  in period  $t$  ( $t = 2, \dots, T_i$ ;  $i \in \Omega_2$ ).

The model for individual earnings is,

$$y_{i1} = \tilde{\beta}'\mathbf{x}_{i1} + \varepsilon_{i1},$$

$$y_{it} = \gamma y_{i,t-1} + (1 - \gamma)\beta'\mathbf{x}_{it} + (1 - \gamma)\tau_i + (1 - \gamma)\phi\varepsilon_{i1} + \varepsilon_{it} \quad (t = 2, \dots, T),$$

$$\varepsilon_{i2} = \tilde{\rho}\varepsilon_{i1} + \eta_{i2},$$

$$\varepsilon_{it} = \rho\varepsilon_{i,t-1} + \eta_{it} \quad (t = 3, \dots, T),$$

$$\varepsilon_{i1} = \alpha_{i1} + h_{i1}^{-1/2}\zeta_{i1},$$

$$\eta_{it} = \alpha_{it} + h_{it}^{-1/2}\zeta_{it} \quad (t = 2, \dots, T),$$

$$\tau_i = \sigma_\tau\zeta_{i0}.$$

$$\zeta_{it} \quad (t = 0, \dots, T; i = 1, \dots, n) \stackrel{iid}{\sim} N(0, 1).$$



The parameters  $\{a_{it}, h_{it}\}$  arise from a discrete mixture of normals model. In this model, the disturbance  $\varepsilon_{it}$  is a discrete mixture of three normals, with one set of parameters; the distribution of  $\eta_{it}$  is a discrete mixture of three normals with another set of parameters.

The prior distribution is composed of the following independent components:

$$\begin{aligned} \delta &= (\tilde{\beta}', \beta')' \sim N(\underline{\delta}, \underline{H}_\delta^{-1}), \quad \underline{H}_\delta \text{ p.d.}, \\ \pi &= (\tilde{\rho}, \rho)' \sim N(\underline{\pi}, \underline{H}_\pi^{-1}), \quad \underline{H}_\pi \text{ p.d.}, \quad -1 < \rho < 1, \\ \phi &\sim N(\underline{\phi}, h_\phi^{-1}), \\ \gamma &\sim N(\underline{\gamma}, h_\gamma^{-1}), \quad -1 < \gamma < 1, \\ \underline{v}_\tau \underline{s}_\tau^2 / \sigma_\tau^2 &\sim \chi^2(\underline{v}_\tau), \quad \underline{s}_\tau^2 > 0, \quad \underline{v}_\tau > 0. \end{aligned}$$

The joint posterior density of the parameters and the latent variables  $(\tau_i)_{i=1}^n$ , conditional on the parameters of the normal mixture process, is

$$\begin{aligned} & p(\tilde{\beta}, \beta, \tilde{\rho}, \rho, \phi, \gamma, \sigma_\tau^2, (\tau_i)_{i=1}^n \mid \{(y_{it}, \mathbf{x}_{it}, \alpha_{it}, h_{it})_{t=1}^{T_i}\}_{i=1}^n, \underline{\delta}, \underline{H}_\delta, \underline{\pi}, \underline{H}_\pi, \underline{\phi}, h_\phi, \gamma, h_\gamma, \underline{s}_\tau^2, \underline{v}_\tau) \\ & \propto \prod_{i=1}^n [p(\{y_{it}\}_{t=1}^{T_i} \mid \{\mathbf{x}_{it}, \alpha_{it}, h_{it}\}_{t=1}^{T_i}, \tilde{\beta}, \beta, \tilde{\rho}, \rho, \phi, \gamma, \sigma_\tau^2, (\tau_i)_{i=1}^n) p(\tau_i \mid \sigma_\tau^2) \\ & \quad p(\delta, \pi, \phi, \gamma, \sigma_\tau \mid \underline{\delta}, \underline{H}_\delta, \underline{\pi}, \underline{H}_\pi, \underline{\phi}, h_\phi, r, h_\gamma, \underline{v}_\tau, \underline{s}_\tau^2) \\ & \quad \propto \prod_{i=1}^n h_{i1}^{1/2} \exp\left[-\frac{1}{2} \sum_{t=1}^n h_{i1} (y_{i1} - \alpha_{i1} - \tilde{\beta}' \mathbf{x}_{i1})^2\right] \\ & \quad \times \prod_{i \in \Omega_2} h_{i2}^{1/2} \exp\left\{-\frac{1}{2} \sum_{i \in \Omega_2} h_{i2} [y_{i2} - \alpha_{i2} - \gamma y_{i1} - \gamma_1 \beta' \mathbf{x}_{i2} - \gamma_1 \tau_i \right. \\ & \quad \left. - (\tilde{\rho} + \gamma_1 \phi)(y_{i1} - \tilde{\beta}' \mathbf{x}_{i1})]^2\right\} \\ & \quad \times \prod_{i \in \Omega_3} \prod_{t=3}^{T_i} h_{it}^{1/2} \exp\left\{-\frac{1}{2} \sum_{i \in \Omega_3} \sum_{t=3}^{T_i} h_{it} [y_{it} - \alpha_{it} - \gamma y_{i,t-1} - \gamma_1 \beta' \mathbf{x}_{it} - \gamma_1 \tau_i \right. \\ & \quad \left. - \gamma_1 \phi(y_{i1} - \tilde{\beta}' \mathbf{x}_{i1}) - \rho(y_{i,t-1} - \gamma y_{i,t-2} - \gamma_1 \beta' \mathbf{x}_{i,t-1} - \gamma_1 \tau_i \right. \\ & \quad \left. - \gamma_1 \phi(y_{i1} - \tilde{\beta}' \mathbf{x}_{i1}))]^2\right\} \\ & \quad \times \sigma_\tau^{-n} \exp\left(-\frac{1}{2} \sigma_\tau^{-2} \sum_{i=1}^n \tau_i^2\right) \exp\left[-\frac{1}{2} (\delta - \underline{\delta})' \underline{H}_\delta (\delta - \underline{\delta})\right] \\ & \quad \times \exp\left[-\frac{1}{2} (\pi - \underline{\pi})' \underline{H}_\pi (\pi - \underline{\pi})\right] \chi_{(-1,1)}(\rho) \exp\left[-\frac{1}{2} h_\phi (\phi - \underline{\phi})^2\right] \\ & \quad \times \exp\left[-\frac{1}{2} h_\gamma (\gamma - \underline{\gamma})^2\right] \chi_{(-1,1)}(\gamma) \sigma_\tau^{-\underline{v}_\tau} \exp\left(-\frac{1}{2} \underline{v}_\tau \underline{s}_\tau^2 \sigma_\tau^{-2}\right). \end{aligned}$$

**Appendix B. Posterior distribution for the marital status model**

The panel consists of  $n$  individuals,  $i = 1, \dots, n$ . In this appendix it is assumed that individual  $i$  is observed in periods  $1, \dots, T_i$ ;  $T_i \leq T$  ( $i = 1, \dots, n$ ). Let  $\Omega_j = \{i: T_i \geq j\}$ , the set of individuals observed in period  $j$ , and let  $N_j$  denote the cardinality of  $\Omega_j$ . The total number of observations is  $N = \sum_{i=1}^n T_i = \sum_{t=1}^T N_t$ . For  $(t = 1, \dots, T; i = 1, \dots, n)$  let

- $d_{it} = 1$  if individual  $i$  is married in period  $t$  and  $d_{it} = 0$  if not,
- $s_{1t} = (p_1 \times 1)$  vector of deterministic variables,
- $s_{it} = (p_2 \times 1)$  vector of deterministic variables and lagged earnings, possibly interacted with  $d_{i,t-1}$  ( $t = 2, \dots, T; i \in \Omega_2$ ),
- $m_{it}^*$  = a latent variable that determines  $d_{it}$ .

The model for marital status is

$$m_{i1}^* = \tilde{\theta}'s_{i1} + \xi_{i1},$$

$$\xi_{i1} \stackrel{iid}{\sim} N[0, (1 - \lambda^2)^{-1}],$$

$$m_{it}^* = \theta's_{it} + \xi_{it} \quad (t = 2, \dots, T),$$

$$\xi_{it} = \lambda \xi_{i,t-1} + \Psi_{it} \quad (t = 2, \dots, T),$$

$$\Psi_{it} \stackrel{iid}{\sim} N(0, 1) \quad (t = 2, \dots, T),$$

$$d_{it} = \begin{cases} 1 & \text{if } m_{it}^* \geq 0, \\ 0 & \text{if } m_{it}^* < 0. \end{cases}$$

The parameters of the model are  $\tilde{\theta}$ ,  $\theta$  and  $\lambda$ . The prior distribution has two independent components,

$$\mu = (\tilde{\theta}', \theta') \sim N(\underline{\mu}, \underline{H}_\mu^{-1}), \quad \underline{H}_\mu \text{ p.d.},$$

$$\lambda \sim N(\underline{\lambda}, \underline{h}_\lambda^{-1}), \quad \underline{h}_\lambda > 0, \quad -1 < \lambda < 1.$$

The joint posterior density of the parameters and latent variables is

$$p(\tilde{\theta}, \theta, \lambda, \{(m_{it}^*)_{i=1}^T\}_{i=1}^n \mid \{(d_{it}, s_{it})_{i=1}^T\}_{i=1}^n, \underline{\mu}, \underline{H}_\mu, \underline{\lambda}, \underline{h}_\lambda)$$

$$\propto (1 - \lambda^2)^{n/2} \exp \left[ -\frac{1}{2} (1 - \lambda^2) \sum_{i=1}^n (m_{i1}^* - \tilde{\theta}'s_{i1})^2 \right]$$

$$\begin{aligned}
 & \times \exp \left\{ -\frac{1}{2} \sum_{i \in \Omega_2} [m_{i2}^* - \theta' s_{i2} - \lambda(m_{i1}^* - \tilde{\theta}' s_{i1})]^2 \right\} \\
 & \times \exp \left\{ -\frac{1}{2} \sum_{i \in \Omega_3} \sum_{t=3}^{T_i} [m_{it}^* - \theta' s_{it} - \lambda(m_{i,t-1}^* - \tilde{\theta}' s_{i,t-1})]^2 \right\} \\
 & \times \prod_{i=1}^n \prod_{t=1}^{T_i} I_{S(d_n)}(m_{it}^*) \\
 & \times \exp \left[ -\frac{1}{2} (\mu - \underline{\mu})' \underline{H}_\theta (\mu - \underline{\mu}) \right] \\
 & \times \exp \left[ -\frac{1}{2} h_\lambda (\lambda - \underline{\lambda})^2 \right] \cdot \chi_{(-1,1)}(\lambda), \tag{B.1}
 \end{aligned}$$

where  $S(0) = (-\infty, 0)$  and  $S(1) = [0, \infty)$  in (B.1).

### Appendix C. Prior distributions

All prior distributions are proper. In setting the priors, our guiding principle is that the priors should be flexible to make very large but reasonable effects likely.

#### C.1. Priors in the earnings model

In our model there are 10 covariates for the log of real earnings in the first period ( $k_1 = 10$ ) and 16 covariates for the log of real earnings in the other periods ( $k_2 = 16$ ). The prior distribution for the 26 corresponding coefficients  $\delta' = (\beta', \beta')$  is multivariate normal.

The first 8 covariates are the same for all periods. The covariates and their prior means and standard deviations are given in Table 12.

Prior means reflected our best subjective judgement of the most plausible effects of each of the covariates on the log of real earnings. Prior standard deviations were chosen to reflect considerable uncertainty, and to be consistent with our role as investigators rather than clients. In the case of the missing education dummies, we had very little idea what the effects might be, since we have little insight into the missing variable process.

The remaining covariates pertain to the effects of age and education on expectations of log earnings. Our prior for these coefficients is independent of the prior distribution of the eight covariates just discussed. It is constructed by thinking about expected log earnings at eight combinations of age and education: ages 25, 35, 45, and 55; and 12 and 16 yr of education. In our model, the only covariates that affect expected earnings *growth* are age and education. Denote the impact of age and education on expected log earnings growth from

Table 12

Covariate	Prior mean	Prior stan. dev.
Father education missing dummy (1 if missing, 0 if present)	0.05	0.20
Father high school dummy (1 if known high school grad, else 0)	0.05	0.10
Father college dummy (1 if known college grad, else 0)	0.10	0.10
Mother education missing dummy (1 if missing, 0 if present)	0.05	0.20
Mother high school dummy (1 if known high school grad, else 0)	0.05	0.10
Mother college dummy (1 if known college grad, else 0)	0.10	0.10
Race dummy (1 if nonwhite, 0 if white)	− 0.10	0.10
Lagged marital status dummy (1 if married, 0 if not married)	0.20	0.20

period 1 to period 2 by  $G(a_1, a_2, e_1, e_2)$ . There are seven independent prior distributions for values of this function, with means and standard deviations given in Table 13.

Thus, real earnings of a high school graduate with no college are expected to grow at a 1.5% annual rate from 25 to 35, at 1% from 35 to 45, and at 0.5% from 45 to 55. The earnings of a college graduate are expected to be 1.3 times those of a high school graduate at age 25, 1.4 times at 35, 1.45 times at 45, and 1.5 times at 55. Prior standard deviations are sufficiently large to make other reasonable growth rates – including none at all – plausible.

We complete our prior distribution for the effects of age and education on expected log earnings with an independent distribution for expected log earnings of a white high school graduate for whom it is known that neither parent graduated from high school. This distribution has mean 8 and standard deviation 4, which is quite diffuse and includes unreasonable as well as reasonable values of expected log earnings.

For periods beyond the first, the eight age and education covariates are the full interaction of age raised to the powers 0, 1, 2, and 3, on the one hand, with education raised to the powers 0 and 1, on the other. The foregoing prior distribution for eight age–education combinations induces the prior distribution on the coefficients corresponding to these covariates through a standard linear transformation. For the first period the covariates are the intercept and age. The prior distribution on the level and  $G(25, 25; 12, 16)$  provide the prior distribution on the corresponding coefficients.

Table 13

Age/education linear combination	Prior mean	Prior stan. dev.
G (25, 35; 12, 12)	0.15	0.10
G (35, 45; 12, 12)	0.10	0.10
G (45, 55; 12, 12)	0.05	0.10
G (25, 25; 12, 16)	0.26	0.15
G (35, 35; 12, 16)	0.34	0.20
G (45, 45; 12, 16)	0.37	0.225
G (55, 55; 12, 16)	0.40	0.25

Priors for other parameters in the earnings model were set as follows.

$\gamma, \rho_1, \rho_2$ : Normal priors for these serial persistence parameters were each centered at 0.5, with a standard deviation of 0.5. As explained in Appendix A, the priors are then truncated to force the parameters to lie between  $-1$  and  $1$ . These priors were mutually independent and independent of priors for all other parameters.

$\phi$ : The value 0.4 was taken to be a large persistent effect of the first-year shock. Since we are not very certain about this, prior standard deviation is set to 0.5. This prior is independent of priors for all other parameters.

$\sigma_\tau^2$ : If heterogeneity not associated with the covariates contributed typically to variations by a factor of two in earnings, then these terms would be about  $[\log(2)]^2 = 0.480$ . We assigned the prior  $0.4/\sigma_\tau^2 \sim \chi^2(1)$ , implying that *a priori*  $P(0.104 \leq \sigma_\tau^2 \leq 101.7) = 0.9$ .

### C.1.1. Shock distributions

The disturbances  $\varepsilon_{i1}$  and  $\eta_{it}$  ( $t \geq 2$ ) are each modeled as mixtures of three normal distributions. The prior distribution of the parameters of these distributions are the same for each kind of disturbance. The prior distribution was developed after inspection of residuals in models with i.i.d. normal disturbances. Using ‘ $x_i$ ’ do denote a generic random variable, the three-normal mixture distribution can be described

$$x_i \sim N(\alpha_{(j)}, h_{(j)}^{-1}) \quad \text{with probability } p_j \quad (j = 1, 2, 3).$$

(For a complete discussion, see Appendix F.) We impose the identifying restriction  $\alpha_{(1)} \leq \alpha_{(2)} \leq \alpha_{(3)}$ .

Subject to this restriction, the prior distribution for  $\alpha' = (\alpha_{(1)}, \alpha_{(2)}, \alpha_{(3)})$  is

$$\alpha_{(1)} \sim N(-3, 1)$$

$$\alpha_{(2)} = 0$$

$$\alpha_{(3)} \sim N(0.1, 0.1)$$

(The restriction  $\alpha_{(2)} = 0$  is a normalization, since an intercept term,  $\tilde{\beta}_9$  or  $\beta_9$  is already included as a covariate.)

The independent prior distributions for the  $h_{(j)}$  are

$$200h_{(1)} \sim \chi^2(100) \Rightarrow P[1.268 \leq h_{(1)}^{-1/2} \leq 1.602] = 0.9;$$

$$125h_{(2)} \sim \chi^2(100) \Rightarrow P[1.002 \leq h_{(2)}^{-1/2} \leq 1.266] = 0.9;$$

$$10h_{(3)} \sim \chi^2(100) \Rightarrow P[0.284 \leq h_{(3)}^{-1/2} \leq 0.358] = 0.9.$$

The prior distribution for  $\mathbf{p}$  is multivariate beta, with parameters 2, 48, and 50, implying modal values of 0.0103, 0.4895, and 0.5052 for the prior distribution of  $\mathbf{p}$ . The marginal prior distributions are univariate beta, with

$$P(0.0036 \leq p_1 \leq 0.0470) = 0.90,$$

$$P(0.3984 \leq p_2 \leq 0.5621) = 0.90,$$

$$P(0.4181 \leq p_3 \leq 0.5819) = 0.90.$$

### C.2. Priors in the marital status model

There are 3 covariates for the log of real earnings in the first period ( $p_1 = 3$ ) and 9 covariates for the log of real earnings in the other periods ( $p_2 = 9$ ). The prior distribution for the 12 corresponding coefficients  $\mu' = (\theta', \theta')$  is multivariate normal.

Since this is a probit model, scale effects are determined by the standard normal distribution for the innovations in the marital status state equation. In thinking about the magnitude of any covariate coefficient, it is necessary to consider the effect of a change in the covariate on the marital status from a particular starting probability.

Of the 9 covariates for periods beyond the first, 6 correspond to interactions of education and age: the full interaction of age raised to the powers 0, 1, and 2, on the one hand, with education raised to the powers 0 and 1, on the other. Our prior distribution for the corresponding coefficients is independent of the prior distribution for the coefficients of the other three covariates. It is constructed by thinking about plausible effects of different age and education on marital status probability, when that probability is about 0.5. Let  $\Delta(a_1, a_2, e_1, e_2)$  denote the impact on the *probit* of a change in age from  $a_1$  to  $a_2$  and a change in education from  $e_1$  to  $e_2$ . We adopt independent priors for five linear combinations of the coefficients and they are given in Table 14.

A sixth prior distribution, independent of these five and the other coefficients, assigns a mean of  $\Phi^{-1}(0.5) = 0$  and standard deviation of  $[\Phi^{-1}(0.6) - \Phi^{-1}(0.5)]/2 = 0.255$  to the probability of being married at age 25 and education 12.

Of the remaining three covariates for periods beyond the first, the nonwhite dummy is assigned a mean of 0 and a standard deviation of  $\Phi^{-1}(0.6) - \Phi^{-1}(0.5) = 0.255$ . Lagged marital status is given both a mean and standard

Table 14

Age/education linear combination	Prior mean	Prior stan. dev.
$\Delta$ (25, 40; 12, 12)	0	$\Phi^{-1}(0.6) - \Phi^{-1}(0.5) = 0.255$
$\Delta$ (40, 55; 12, 12)	0	$\Phi^{-1}(0.6) - \Phi^{-1}(0.5) = 0.255$
$\Delta$ (25, 25; 12, 16)	0	$\Phi^{-1}(0.6) - \Phi^{-1}(0.5) = 0.255$
$\Delta$ (40, 40; 12, 16)	0	$\Phi^{-1}(0.6) - \Phi^{-1}(0.5) = 0.255$
$\Delta$ (55, 55; 12, 16)	0	$\Phi^{-1}(0.6) - \Phi^{-1}(0.5) = 0.255$

deviation of  $\Phi^{-1}(0.8) - \Phi^{-1}(0.2) = 0.68$ . The mean for the coefficient on lagged earnings is 0. The standard deviation is set by reasoning that an increase in log earnings of one unit might change marital status probability by 0.2 when that probability was about 0.5 to start: so the standard deviation was chosen to be  $\Phi^{-1}(0.6) - \Phi^{-1}(0.4) = 0.510$ . In choosing all of these values, the intention is to select reasonable means, and standard deviations sufficiently large as to make all plausible values fall within two standard deviations.

In the first period the only covariates are intercept, education, and the race dummy. The prior distributions for the race dummy coefficient is the same as in the later periods, and the prior distribution for intercept and age is inferred from the distributions above.

The prior specification of the model is completed by assigning a normal distribution to the serial correlation coefficient  $\lambda$ , with a mean and standard deviation of 0.7.

## References

- Chib, S., Greenberg, E., 1995. Understanding the Metropolis-Hastings algorithm. *The American Statistician* 49, 327–335.
- Chib, S., Greenberg, E., 1996. Markov chain Monte Carlo simulation methods in econometrics. *Econometric Theory* 12, 409–431.
- Escobar, M.D., West, M., 1995. Bayesian density estimation and inference using mixtures. *Journal of the American Statistical Association* 90, 577–588.
- Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1995. *Bayesian Data Analysis*. Chapman and Hall, London.
- Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. *Statistical Science* 7, 457–472.
- Geweke, J., 1989. Bayesian inference in econometric models using Monte Carlo integration. *Econometrica* 57, 1317–1339.
- Geweke, J., 1996. Posterior Simulators in Econometrics. In: Kreps, D., Wallis, K.F. (Eds.), *Advances in Economics and Econometrics: Theory and Applications*, Vol. III. Cambridge University Press, Cambridge, 1997, pp. 128–165.
- Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1996. *Markov chain Monte Carlo in practice*. Chapman and Hall, London.
- Gottschalk, P., 1982. Earnings mobility: permanent change or transitory fluctuations? *Review of Economics and Statistics* 64, 450–456.

- Gottschalk, P., Moffit, R., 1994. Trends in the autocovariance structure of earnings in the U.S.: 1969–1987. Working Paper, Department of Economics, Brown University.
- Hirano, K., 1998. Semiparametric Bayesian models for dynamic earnings data. Harvard University Working Paper (March 3).
- Horowitz, J.L., Markatou, M., 1996. Semiparametric estimation of regression models for panel data. *Review of Economic Studies* 63, 145–168.
- Kiefer, J., Wolfowitz, J., 1956. Consistency of the maximum likelihood estimator in the presence of infinitely many nuisance parameters. *Annals of Mathematical Statistics* 27, 887–906.
- Lillard, L., Willis, R., 1978. Dynamic aspects of earnings mobility. *Econometrica* 46, 985–1012.
- MacEachern, S., Mueller, P., 1995. Estimating Mixture of Dirichlet Process Models. Ohio State University, Technical Report.
- MaCurdy, T., 1982. The use of time series processes to model the error structure of earnings in a longitudinal data analysis. *Journal of Econometrics* 18, 83–114.
- McCall, J.J., 1973. *Income Mobility, Racial Discrimination, and Economic Growth*. Lexington Books, Lexington, MA.
- Redner, R.A., Walker, H.F., 1984. Mixture densities, maximum likelihood and the EM algorithm. *SIAM Review* 26, 195–239.
- Roeder, K., Wasserman, L., 1997. Practical Bayesian density estimation using mixtures of normals. *Journal of the American Statistical Association* 92, 894–902.
- Shorrocks, A., 1976. Income mobility and the Markov assumption. *Economic Journal* 86, 566–578.
- Tierney, L., 1994. Markov chains for exploring posterior distributions (with discussion and rejoinder). *Annals of Statistics* 22, 1701–1762.
- Tittertoning, D.M., Smith, A.F.M., Makov, U.E., 1985. *Statistical Analysis of Finite Mixture Distributions*. Wiley, Chichester.
- West, M., Harrison, J., 1989. *Bayesian Forecasting and Dynamic Models*. Springer, New York.